Agenda and Notes

CC CoP Webinar
February 22, 2021, 1-3 pm CT
Hangouts Meet
Join online: meet.google.com/jmq-hrrr-rbc
Phone:+1 929-252-0881, PIN: 989 421 120#

Meeting Purpose and Objectives:

- Provide an overview of different cloud computing/storage opportunities for the community to take advantage of
- Hold a discussion with all participants centered around water model data (earth systems data) in the cloud (generating it, accessing it, using it downstream).
- Discuss future efforts and activities with the audience

Outline

1. Introduction and set the stage
2. Patrick Tripp (RPS) IOOS Cloud Sandbox
3. Rich Signell (USGS) Open Architectures for Cloud-Native Earth System Analytics
4. Jena Kent (NOAA/CO-OPS) NOAA's Big Data Program
5. Dan Morris (Microsoft) Microsoft AI for Earth
6. Future engagement opportunities and adjourn

Documents
CoP presentation
Rich presentation
Dan presentation

Attendees
Cayla Dean, Brenna Sweetman, Patrick Tripp, Rich Signell, Jena Kent, Dan Morris, Adrienne Simonson, Aijun Zhang, Alex Prusevich, Alison Macneil, Amin Kiaghadi, Camaron George, Cheryl Ann Blain, Chris Massey, Chris Paternostro, Christina Urizar, Debra Hernandez, Degui Cao, Derrick Snowden, Dina Sang, Evan Turner, Frank Tsai, Gina Martinez, Graeme Aggett, Hamed Moftakhari, Hassan Mashriqui, Ilya Rivin, Jeff Arnold, Jennifer Arrigo, Jennifer McGee, Jenny Dissen, Jim McManus, Jingtao Xu, John Schmidt, John Warner, John Wilkin, Jonathan O'Neil, Joseph Zhang, JS Allen, Julio Zyserman, Jungwoo Lee, Katherine Powell, Katie Kirk, Kendra Dresback, Kevin De Santiago, Lei Shi, Lianyuan Zheng, Linlin Cui, Liv Herdman, Marouane Temimi, Melissa Lupher, Nels Frazier, Otis Brown, Pat Burke, Patrick Keown, Paul Bradley, Ram Neupane, Renee Collini, Rick Luettich, Ruoying He, Sadiq Khan, Shahidul Islam, Tarandeep Kalra, Thomas Williams, Tom Shyka

Notes:
Questions for Patrick Tripp:

Thomas Williams1:19 PM
so you're using Prefect to submit an AWS Batch job?

Chris Massey1:22 PM
What has been your experience on wait times for cloud computing resources to be spun-up for use? Do larger resources request (50 nodes or more) require any special handling?

Agenda and Notes

Amin Kiaghadi1:23 PM
how is this different from DesignSafe? (https://www.designsafe-ci.org/)

Questions for Richard Signell

Patrick Tripp1:41 PM
I noticed it's a requester pay bucket and am wondering what those monthly costs are to the requester.

Questions for Jena Kent

Derrick Snowden - NOAA Federal2:02 PM
How does NOAA prioritize which data CISESS will move from NOAA to BDP/Cloud?

Thomas Williams2:04 PM
Thanks for making this data available via the big cloud providers - it definitely makes our lives easier. Are there any efforts to make data available in systems like Google BigQuery in addition to raw NetCDF files?

Adrienne Simonson - NOAA Federal2:07 PM
Each LO has an internal process to determine dataset priorities. If you are interested, contact your EDMC representative.

Brian Blanton2:10 PM
Perhaps this was answered by Rich earlier, but is there a cloud-aware THREDDS server

Questions for Dan Morris:

Richard Signell2:23 PM
Is that really true that the demos would go down if we all visited at the same time, or was that a joke?

Richard Signell2:25 PM
Also a curated list of pre-processed climate datasets for ML here:
http://mldata.pangeo.io/preprocessed_datasets.html
Patrick Tripp2:25 PM
very cool AI stuff, thanks. and where did you get the Atari t-shirt?

Thomas Williams2:33 PM
So like this will be like Databricks but Pangeo?

Open discussion:

Derrick Snowden - NOAA Federal2:35 PM
Do you foresee a relationship between NOAA BDP and AI for Earth? Can we assume that BDP will be your source for NOAA data?

Brian Blanton2:41 PM
Since there are several large cloud providers, tools will eventually need to be provider agnostic (except for the billing...)

Agenda and Notes

Dan Morris2:41 PM
@Rich, I honestly have no idea what would happen if everyone visited our demos all at once. Let me know if you find out!

Richard Signell2:42 PM
Indeed Brian, we need to be able to work on any cloud, and work toward minimal configuration and tooling changes to do so

Richard Signell2:44 PM
Dan, if they did crash, I guess that would mean they are have intentionally controlled instance or serverless computation limits?

Derrick Snowden - NOAA Federal2:44 PM
@Dan, do you solicit themes from the community for the AI4Earth workshops? Coastal Coupling on Azure next summer?

Dan Morris2:46 PM
@Derrick, the AI for Earth summits I referred to are pretty specific to getting folks started on the cloud (exactly the stuff we're talking about now), but we also run ~monthly sessions with our grantee community that vary topically. A set of talks/demos/etc. on coastal coupling on the cloud could be great, or we could carve out a session as another community event like ESIP.

Derrick Snowden - NOAA Federal2:49 PM
We need to expand that definition to NOAA partners who do work funded by NOAA
Cayla Dean - NOAA Affiliate2:49 PM
If there is a data set you would like BDP to host, please type it in the chat

Richard Signell2:50 PM
Derrick & Jena, so does a IOOS regional modeler quality to have the data handled by the BDP?

Richard Signell2:50 PM
https://www.nationalgeographic.com/environment/article/biden-commits-to-30-by-2030-conservation-executive-orders

Tom Shyka2:51 PM
NERACOOS is working with BDP to submit the 30 year hindcast product from NECOFS.

Jena Kent - NOAA Federal2:51 PM
Rich, can you reach out to the BDP offline?

Richard Signell2:51 PM
sure