OPeNDAP for Data Providers

Tuesday, July 20 • 11:00am - 12:30pm

More info & slides on Sched Session Page: https://sched.co/jMNz

---

Attendance & Check in

Add your Name / affiliation / pronouns / location / social media to a bullet

- .Bob Simons / NOAA Fisheries ERD / he / Monterey / (not very social)
- .Jolyon Ralph / mindat.org~Hudson Institute of Mineralogy / he him/ London UK / @jolyonralph
- Ed Armstrong/ NASA JPL /Pasadena, CA/
- Adam Kennedy / AND LTER / he,him / OR / @hja_live
- Ted Habermann - Metadata Game Changers
- Aleksandar Jelenak / NASA-EED2-HDF Group / he / Virginia
- Aaron Friesz - NASA LP DAAC
- Xiaohua Pan/ NASA GSFC GES DISC/ She/ Greenbelt, MD, USA
- Ross Bagwell / NASA/SSAI / Maryland
- Armin Mehrabian / NASA GES DISC / he, him/
- Peter MacHarrie /NASA/ASRC, Maryland
- Marina Frants/UCSD/she,her/La Jolla, CA
- Ethan Davis, UCAR Unidata, he/him, Boulder, CO
- Daniel Fuka/Virginia Tech/Albuquerque, NM
- Lisa Kaser, NSIDC, she/her, Boulder, CO
- Candida Dewes, NSIDC, she/her, Westminster, CO
- Rebecca Hudak, R2R, she/her, Maine
- Vincent Inverso, NASA (ESDIS), he/him, Philadelphia
- David Shumway / University of Illinois, Chicago / Chicago, IL
- Ned Molder / ESIP Fellow, UW-Madison Nelson Institute / he/him / @nedmolder
- Stephen Haddad / Uk Met Office / he/him / Exeter, UK / @stevehadd
- Shawn Smith / COAPS, Florida State University / Tallahassee / srsmith@fsu.edu
- Chris Lynnes / NASA
- Joe Lee / The HDF Group - NASA EED2 / he / Champaign, IL / @hdfeos
- Vitor S. Martins, Michigan State Univ., MI
- Alan Yang / UCSD/ he him/La Jolla, CA
- Ellen Johnson / MathWorks
- Martin Weinelt / Helmholtz Metadata Collaboration - HMC / Kiel, Germany / mweinelt@geomar.de
- Emery Boose, Harvard Forest LTER

2021 ESIP Summer Meeting (July 19th-23rd, 2021)

https://2021esipsummermeeting.sched.com/

Meeting Takeaways Doc

Community Participation Guidelines & Reporting: Click HERE

- Bruce Wilson, he/him, ORNL DAAC
- Patrick Quinn / NASA EED-2 Element 84 / he / Pennsylvania / patrick@element84.com
- Sean Arms, UCAR/Unidata, sarms@ucar.edu (@lesserwhirls)
- Sheri Phillips, NOAA/NCEI Silver Spring MD
- Steve Hansen, EDGI / he-him / Bellingham, WA
- Binita KC, NASA GSFC/ Greenbelt MD, she/her
- Allan Doyle, NASA ESO, he/him, Boston, MA
- Nga Chung, NASA/JPL/PO.DAAC
- Dave Meyer, NASA/GES DISC
- H. K. "Rama" Ramapriyan SSAI/NASA GSFC (ESDIS)
- Sudhir Shrestha (NOAA Office of Water Predictions)
- Peisheng Zhao, NASA GES DISC
- Micah Wengren, NOAA/NOS/IOOS
- Yaxing Wei, ORNL DAAC
- Nick Doty / NASA - ESDIS / he, him / GSFC
- Long Pham, NASA-GES DISC
- Jason Sharp, NOAA NCEI

Agenda

At session start:

- **Welcome**
    - Orient attendees to tech (Zoom and QiqoChat)
    - Invite attendees to share their name/affiliation/location in the Zoom chat
    - Invite others to take notes in this document (if desired)
    - Assign any other roles needed
- **Community Participation Guidelines Reminder**
    - By participating in this session, you agree to adhere to the ESIP Community Participation Guidelines
    - Report an Issue: https://www.integritycounts.ca/org/esip
- **Goals of this Session**
    - List goals

Resources for this session (link supplementary docs or presentations here)

- .
- .

Before session close:
- **Capture 3 Takeaways & Shared insights**

  - What did we discover?
  - How can we help each other?
  - What knowledge and resources do we have to share?
  - Important remaining questions?

- **Closing circle**
  - Share any reminders
    - Come to our next Cluster Call? on THIS DAY & TIME. Here's what you can expect to see when you show up.
  - Shout-outs & thanks
  - Requests for peer assists

5 ways to stay involved with ESIP
- Join the conversations on our ESIPFed Slack Channel and say hello on #general.
- Keep up with the latest from in and around ESIP by signing up for the ESIP Monday Update or the mailing list of this group.
- Share your feedback, questions, and suggestions with our Twitter community @esipfed and #esipfed.
- Hop on an ESIP Telecon that interests you - no need to RSVP!
- At this meeting, check out the Research Showcase Poster and Demo Gallery.

Feedback
Please take a moment to share your thoughts on this session in Sched by going to the Sched session page linked at the top of this document.

Paste in Zoom Chat notes that you would like to save here

10:32:31 From ellenj to Everyone : Thanks Patrick - this is great.
10:35:32 From Joe Lee / HDF / NASA EED2 to Everyone : I think the term "cloud-optimized" is too broad. I think it should be refined and limited to "AWS S3-optimized."
10:35:44 From ellenj to Everyone : Two questions: 1. Patrick - EOSDIS Zarr store does not utilize the netcdf 4.8.0 (that has Zarr backend)?
10:36:12 From ellenj to Everyone : 2. Anyone -- what is short version of what DMR++ is?

10:37:34 From tedhabermann to Everyone : DMR++ is traditional DMR with chunk locations and sizes.

10:38:00 From Chris Lynnes to Everyone : and DMR itself is the Dataset Metadata Response.

10:38:17 From ellenj to Everyone : thanks Ted and Chris!

10:38:43 From Chris Lynnes to Everyone : I wonder if Athena over S3 might be tractable for Ted's files...

10:39:02 From tedhabermann to Everyone : I agree - very interesting work - thanks Patrick!

10:39:38 From tedhabermann to Everyone : One other question - any role for xArray in your work?

10:41:05 From Patrick Quinn to Everyone : It's compatible with xarray, and that can help a lot in interpreting things like scale/offset, too

10:41:37 From Ethan Davis (Unidata) to Everyone : Thanks Patrick! Great presentation.

10:42:19 From tedhabermann to Everyone : But your ds object was not an xArray dataset - correct?

10:45:19 From Patrick Quinn to Everyone : not an xarray dataset, but xarray is built to work with NetCDF4, which most of our cloud data is, and it can work with zarr stores

10:53:45 From Joe Lee / HDF / NASA EED2 to Everyone : Since ERDDAP is pure Java, you may want to try a new pure Java HDF5 on S3 reader: https://github.com/jamesmudd/jhdf

10:56:12 From Sean Arms (Unidata) to Everyone : netCDF-Java is capable of reading from S3.

11:00:11 From Chris Lynnes to Everyone : @Bob, what size files are we talking about for ERDDAP?

11:00:24 From Aleksandar J. (HDF Group) to Everyone : Exactly my question!

11:05:00 From Ethan Davis (Unidata) to Everyone : Does anyone have experience with accessing non-AWS S3 stores?

11:05:26 From jhrg59701 to Everyone : @erhan - we do have experience with GCS and Azure

11:05:39 From jhrg59701 to Everyone : It's roughly the same

11:05:48 From jhrg59701 to Everyone : DMR++ works with both

11:06:04 From jhrg59701 to Everyone : I think Pangeo has used Zarr with both - at least GCS

11:06:15 From jhrg59701 to Everyone : (GCS == Google Cloud Store)

11:07:08 From Joe Lee / HDF / NASA EED2 to Everyone : https://registry.opendata.aws/terrafusion/ is a good test file on S3 with large file size (20~40G netCDF4).

11:07:54 From Ethan Davis (Unidata) to Everyone : Thanks James. We've worked with an in-house (NCAR) object store through the S3 API as well but don't have a lot of numbers yet.

11:08:47 From ellenj to Everyone : Ethan - about non-AWS data-- do you mean specifically for files backed in Zarr? Or in general? In MATLAB starting in 20b we have support for accessing HDF5 data on S3, Azure, and Hadoop

11:08:48 From Ed Armstrong to Everyone : The MUR SST dataset is also in AWS Open Data Registry

11:12:51 From Ethan Davis (Unidata) to Everyone : Hi Ellen - I was not thinking Zarr specifically, more focused on object stores. Most (all?) commercial cloud providers support the S3 API for

their object stores. The netCDF-Java library uses the AWS SDK for S3 access and can do so on AWS, GCS, an in-house object store, and others (@Sean?).

11:15:06 From Chuan-Yuan Hsu to Everyone : Just a simple question, Can I aggregate the new time step data on a original zarr file?

11:16:35 From Joe Lee / HDF / NASA EED2 to Everyone : @Ethan Davis, is NCAR in-house object store Ceph? (https://docs.ceph.com/en/latest/radosgw/s3/)

11:20:58 From tedhabermann to Everyone : James - do you have the URL for that trade study?

11:21:14 From Sean Arms (Unidata) to Everyone : @Joe Lee, I believe NCAR is using an ActiveScale Object Store, which is probably running a custom fork of Ceph (but not sure).

11:22:11 From Aleksandar J. (HDF Group) to Everyone : @Ted:
https://ntrs.nasa.gov/api/citations/20170009584/downloads/20170009584.pdf

11:22:54 From tedhabermann to Everyone : @aleks - thanks

11:23:46 From Etha