



DEPARTMENT OF ENERGY
**CYBERSECURITY
AND TECHNOLOGY
INNOVATION
CONFERENCE**
May 8-11 | Minneapolis, MN

The Next Generation Earth System Grid Federation



Forrest M. Hoffman (ORNL), Ian Foster (ANL), Sasha Ames (LLNL)
Rachana Ananthakrishnan, Jason Boutte, Nathan Collier, Scott M. Collis, Carlos Downie, Maxwell Grover,
Robert Jacob, Michael Kelleher, Jitendra Kumar, Giri Prakash, Sarat Sreepathi, Min Xu, and Justin Hnilo

DOE Cybersecurity and Technology Innovation Conference

May 9, 2023



U.S. DEPARTMENT OF
ENERGY

Office of
Science



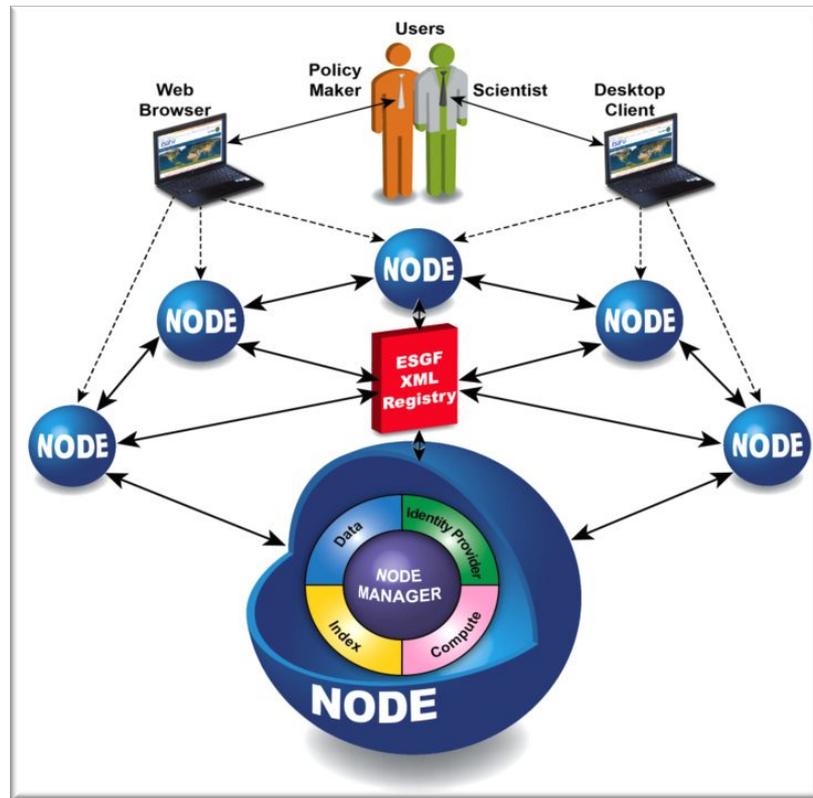
Forrest M. Hoffman, Computational Earth System Scientist

- Group Leader for the ORNL Computational Earth Sciences Group
- 34 years at ORNL in Environmental Sciences Division, then Computer Science and Mathematics Division, and now Computational Sciences and Engineering Division
- Develop and apply Earth system models to study global biogeochemical cycles, including terrestrial & marine carbon cycle
- Investigate methods for reconciling uncertainties in carbon-climate feedbacks through comparison with observations
- Apply artificial intelligence methods (machine learning and data mining) to environmental characterization, simulation, & analysis
- Joint Faculty, University of Tennessee, Knoxville, Department of Civil & Environmental Engineering

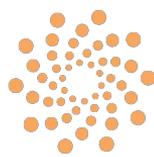


ESGF US 2 What is the Earth System Grid Federation?

- The **Earth System Grid Federation (ESGF)** is an *international consortium* and a *globally distributed peer-to-peer network of data servers* using a common set of protocols and interfaces to archive and distribute climate and Earth system model output and related input, observational, and reanalysis data
- These **Open Science data** are used by scientists all over the world to investigate consequences of possible climate change scenarios and the resulting Earth system feedbacks

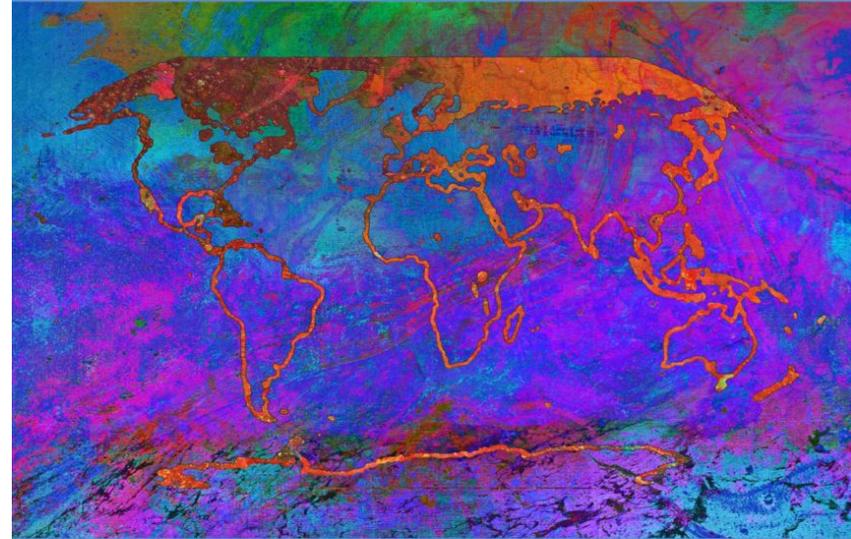


Logos represent primary international contributors: US Department of Energy, NASA, NOAA, NSF, European IS-ENES Project, and Australian NCI



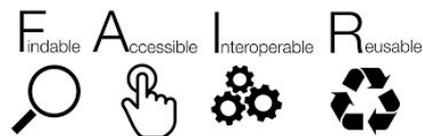
- The United Nations' Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report from Working Group I was released on Monday, August 9, 2021
- All of the climate and Earth system model simulation output underpinning this report was produced by modeling centers participating in the World Climate Research Programme's (WCRP's) sixth phase of the Coupled Model Intercomparison Project (CMIP6)
- Nearly all of that model output was stored in and distributed to researchers via ESGF
- **Data are about the future of life on Earth!**

Climate Change 2021 The Physical Science Basis

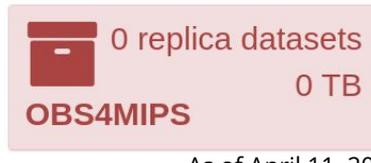
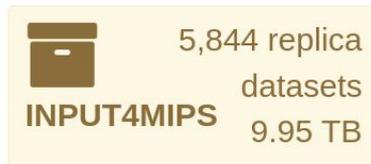
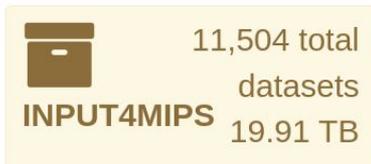
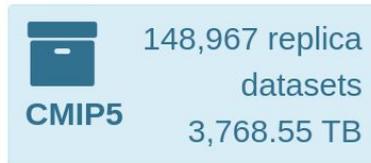
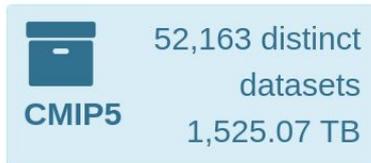
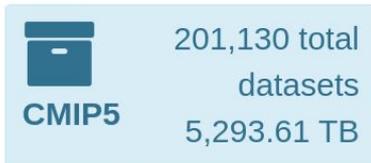
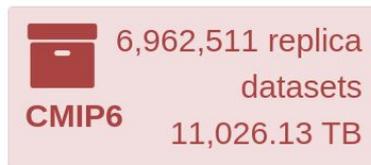
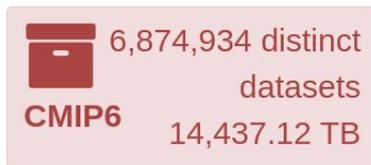
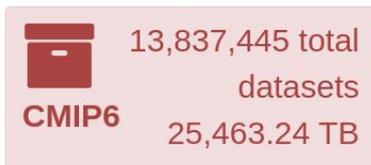




ESGF Holdings are Large



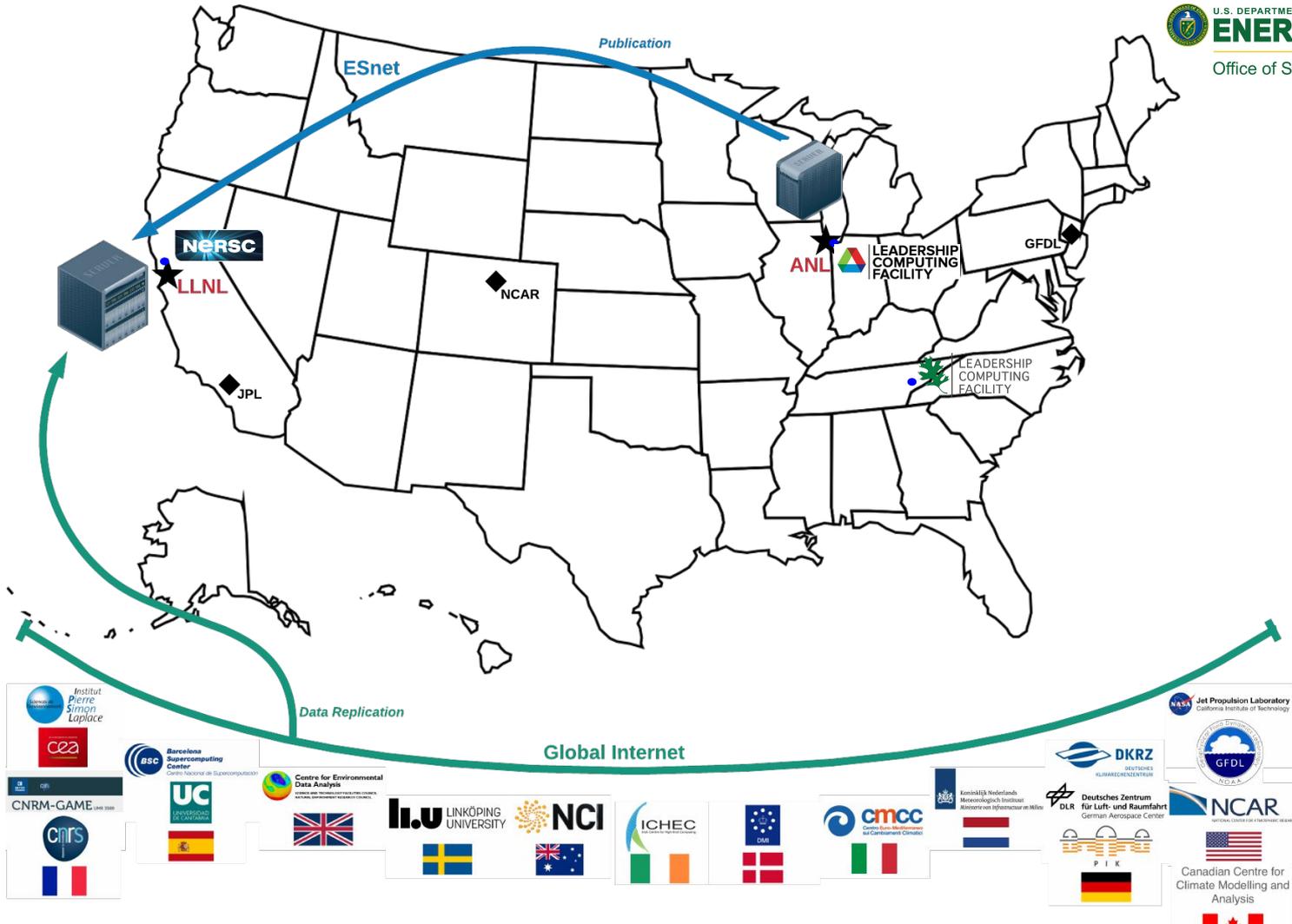
- CMIP5 totals >5 PB (including replicas)
- CMIP6 totals >25 PB (including replicas)
- CMIP7 is expected to have more experiments, high resolution output, and ensembles, totaling ~100 PB
- ESGF is concerned with the full stack security and the integrity of the data, but we are **not** concerned about controlling access to the data





DOE's Current Earth System Grid Federation

- Primary server at LLNL
- LLNL replicates data from the global Federation when possible (primarily up to daily output)
- Independent data node at ANL



ESGF2 US ESnet Global Connectivity

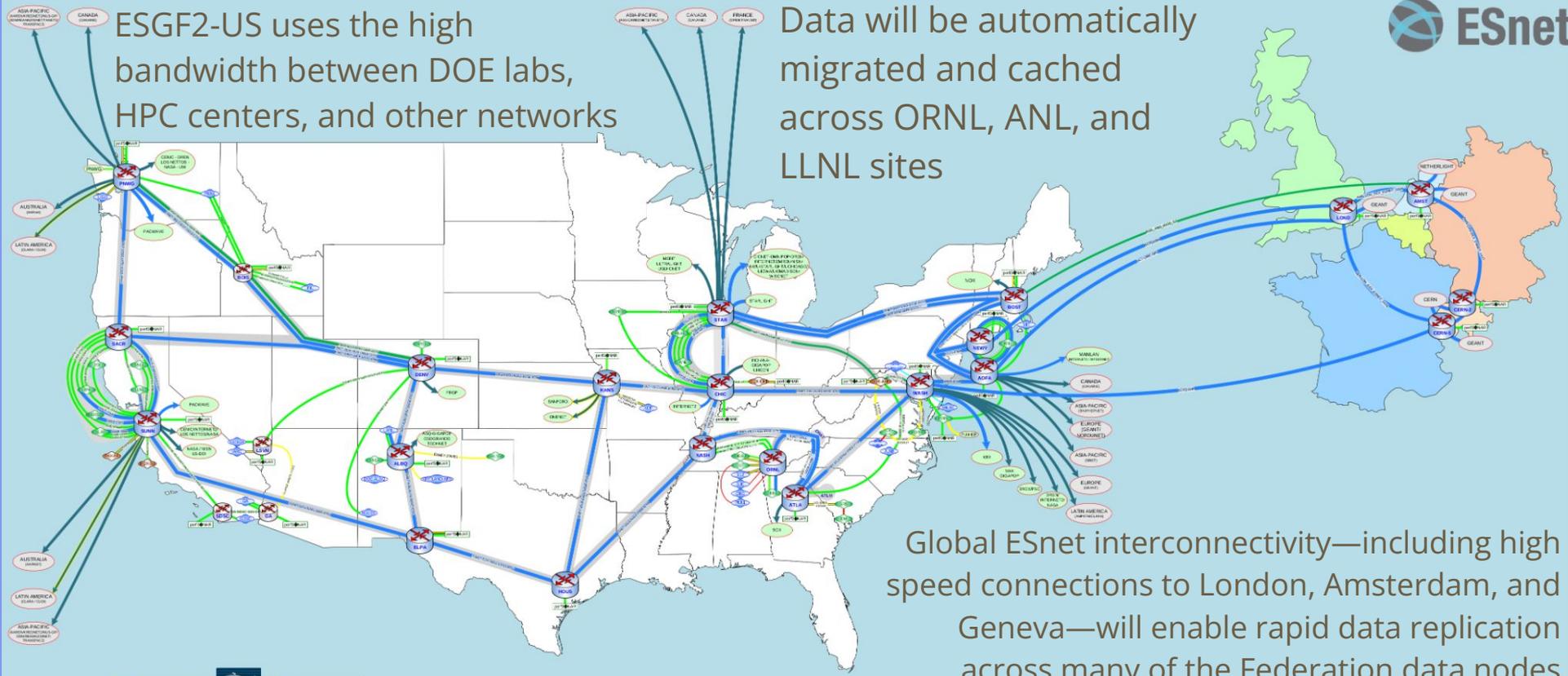


ESnet
ENERGY SCIENCES NETWORK

An ESnet representative is part of the new Resource & Project Liaisons group

ESGF2-US uses the high bandwidth between DOE labs, HPC centers, and other networks

Data will be automatically migrated and cached across ORNL, ANL, and LLNL sites



Global ESnet interconnectivity—including high speed connections to London, Amsterdam, and Geneva—will enable rapid data replication across many of the Federation data nodes

ESGF US 2 Web-based Search Interface

- A new multi-faceted search interface, called **Multigrid**, replaces old interface and adds some requested features
- Requires users to understand simulation experimental design and controlled vocabulary

The screenshot displays the ESGF MetaGrid search interface. The browser address bar shows the URL <https://aims2.llnl.gov/search>. The page header includes the ESGF logo, a project dropdown menu set to 'CMIP6', and a search input field. The main content area shows search results for 'CMIP6' with a total of 12,850,492 results. The results are displayed in a table with columns for 'Dataset Title', 'Files', 'Total Size', 'Version', and 'Download Script'. The table lists several datasets, each with a 'wget' download script and a download icon. The interface also includes a left sidebar for filtering with facets, a top navigation bar with links for Search, Cart, Saved Searches, Node Status, News, Sign In, and Help, and a footer with version information and a 'Made with Hotjar' badge.

ESGF MetaGrid

Search for a keyword

Search

Cart

Saved Searches

Node Status

News

Sign In

Help

Select a Project

CMIP6

CMIP6 Website

Filter with Facets

Expand All

General

histSST-piN2O (2362)

Identifiers

histSST-piNTCF (6445)

histSST-piO3 (1894)

Source ID: historical (592616)

Institution ID: historical-cmp5 (1920)

Source Type: historical-ext (10262)

Experiment ID: historical-with... (567)

Sub-Experiment: land-CCO2 (1563)

Resolutions

Labels

Classifications

Additional Properties

Version Type: Latest

Result Type: Originals and Replicas

Version Date Range: Start ... End d...

Filename

Home / Search

12,850,492 results found for CMIP6

Query String: latest = true

Add Selected to Cart

Save Search

Copy Search

Cart	Dataset Title	Files	Total Size	Version	Download Script
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.Amon.wap.gn	1	572.6 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.clt.gn	1	587.77 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.hfss.gn	1	911.31 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.hfss.gn	1	907.15 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.hursmax.gn	1	764.63 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.hursmin.gn	1	840.16 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.mrso5.gn	1	290.76 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.day.prc.gn	1	363.26 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.Lmon.VegLitter.gn	1	9.6 MB	20190429	wget
<input type="checkbox"/>	CMIP6.ScenarioMIP.CCCma.CanESM5.ssp126.r12i1p2f1.Lmon.grassFrac.gn	1	9.47 MB	20190429	wget

1 2 3 4 5 ... 1285050 > 10 / page

Open as JSON

MetaGrid Version: v1.0.8-beta

Privacy & Legal Notice: <https://www.llnl.gov/disclaimer.html>

Learn about the Department of Energy's Vulnerability Disclosure Program (VDP): <https://d3e.responsibledisclosure.com/hc/en-us>

Overall, what's your experience been with the site?

Please type here...

Made with Hotjar

Next



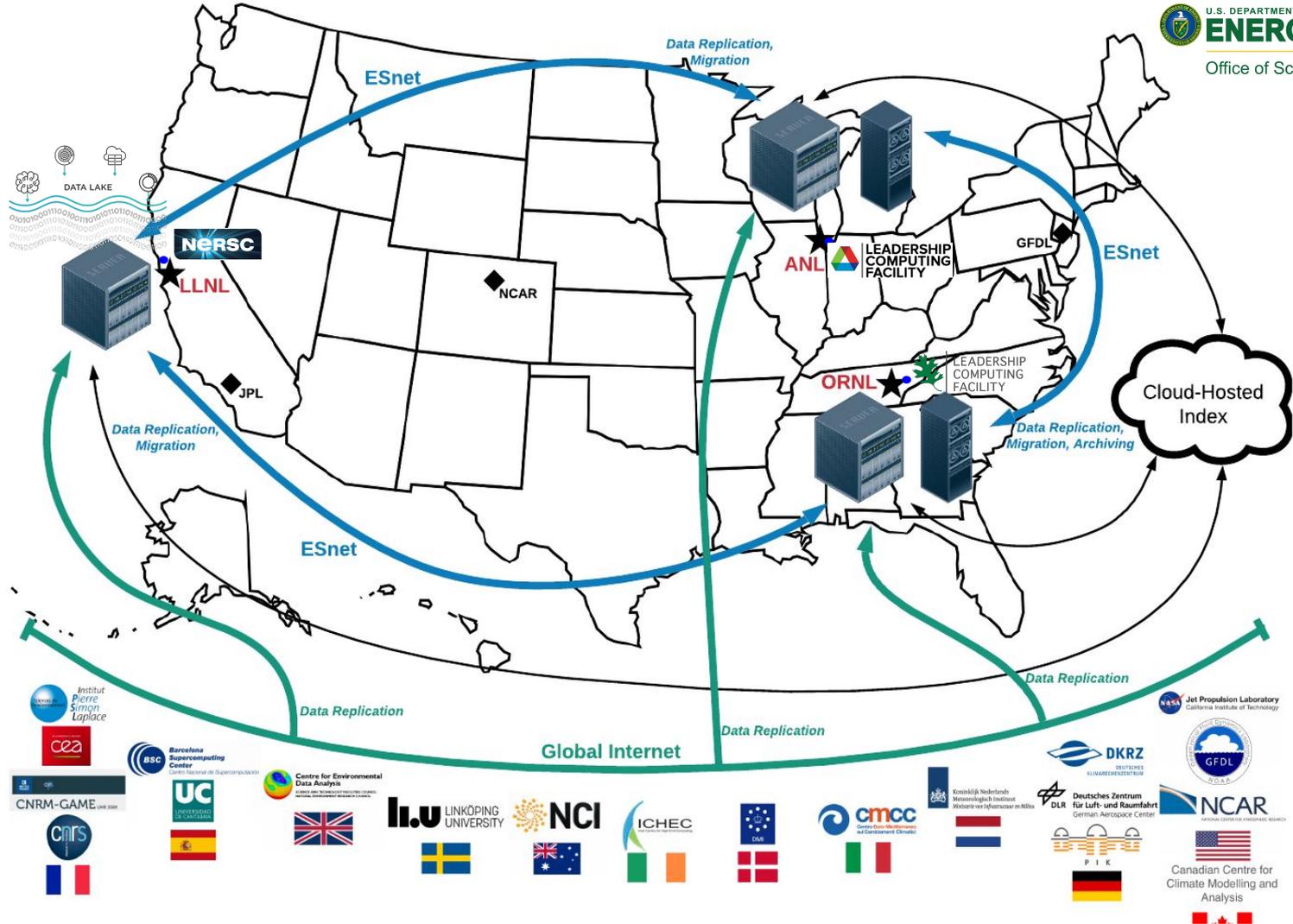
ESGF2 A New Consortium Project in the USA

- New team from **Oak Ridge National Laboratory, Argonne National Laboratory, and Lawrence Livermore National Laboratory** proposed to modernize the data backplane based on the Globus platform
- ESGF2-US proposal was reviewed by panel of 8 scientists in August 2021, and was **selected for funding** by the US Department of Energy starting in FY2022
- In **collaboration with international partners**, the project is developing and will deploy a new architecture based on the *Future Architecture Roadmap*
- In addition, ESGF2-US will develop new **data discovery tools and data access interfaces, server-side computing** (subsetting & summarizing), and **user computing** (Kubernetes & JupyterHub) with improved **user & system metrics**
- ESGF2-US added a **Resource & Project Liaison** group and a **Science, User & Facility Advisory Board** and will offer a help desk/user support



DOE's Next Generation Earth System Grid Federation

- As many as three nodes co-located at DOE's major computing facilities
- Replicating data from the global Federation
- Providing cloud indexing and tape archiving





Design and implementation principles

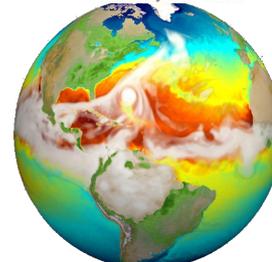
- **Open architecture and protocols**
 - Enable substitution of alternative implementations
- Leverage **highly available and scalable** central services from Globus
 - Reduce complexity, increase reliability, provide economies of scale
- Use proven, modern **security technologies and practices**
 - Integrated access control; protect against attacks and intrusions
- **Use case approach** to design, implementation, and evaluation
 - Ensure that solutions meet real user needs
- Integrated **instrumentation**
 - Metrics drive data management, data access features, capability development
- Focus on **performance** to deal with big data
 - High-speed data transfer, search, server-side processing

ESGF2 US 2 Enabling a new level of research productivity

Logging in with her **institutional credentials**, Samantha is presented with **new data, code, and papers** relevant to her current research. Intrigued by a new report on extreme precipitation events, she examines a **Jupyter notebook** that implements the method used. Wondering how this method would work with higher-resolution E3SM data, she **quickly locates required datasets and runs the notebook on a subset**. Results are promising, so she **shares them with collaborators** via ESGF2-US federated storage, and they agree that a larger ensemble analysis is called for. ESGF2-US confirms that the full ensemble data are available at OLCF, so they submit a request to execute the analysis there. Within 24 hours, **results have been published to ESGF2-US for broader consumption**, along with the notebook used to produce and validate the results.

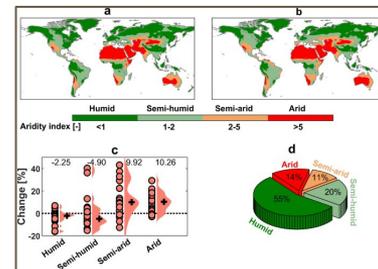


E3SM
Energy Exascale
Earth System Model



OAK RIDGE
National Laboratory

LEADERSHIP
COMPUTING
FACILITY

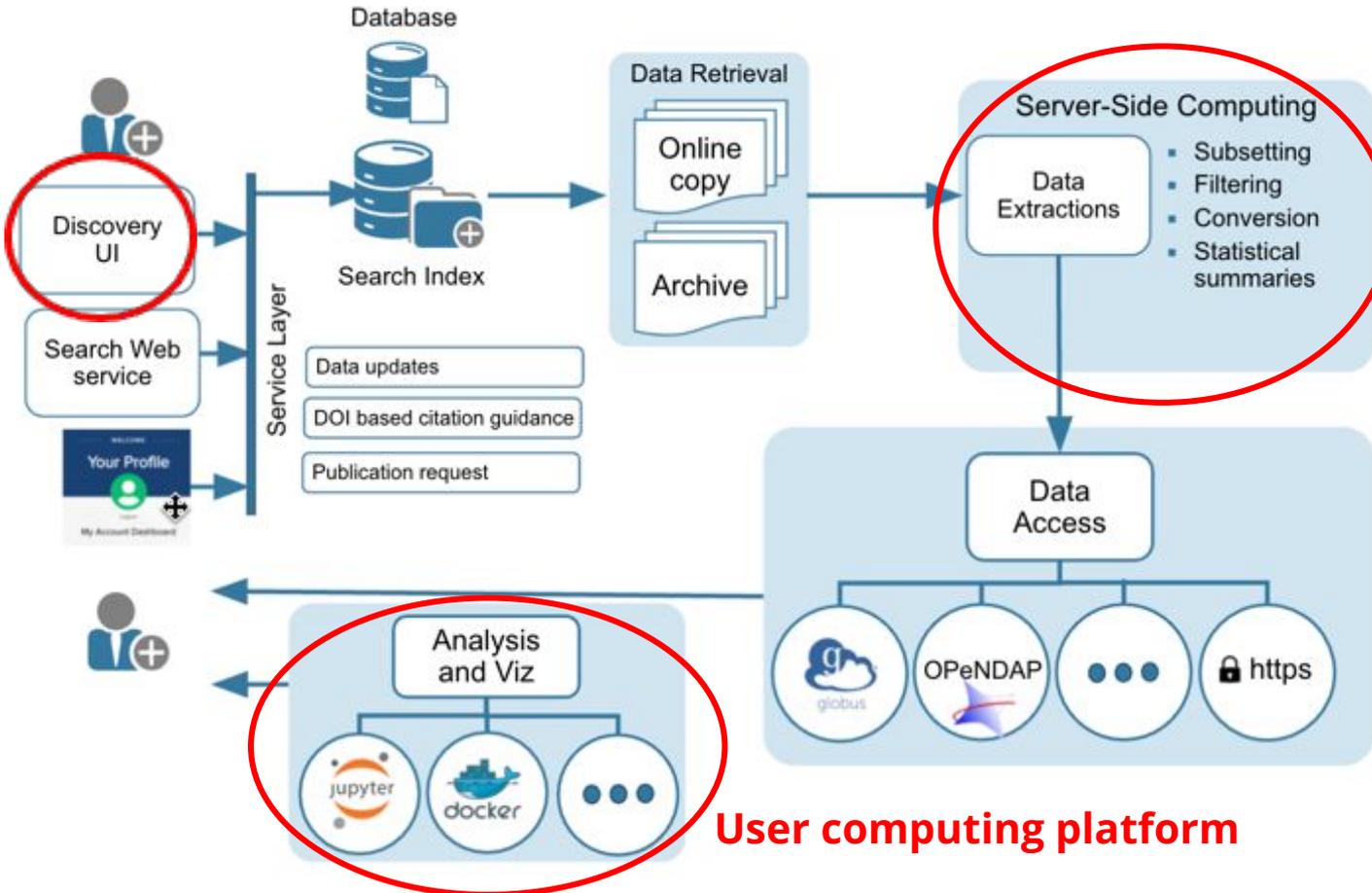


Flood risk increases with water availability



Data Discovery Platform: Architecture

Friendly user interface



Server-side computing platform

User computing platform

Outreach Activities

- Organize **Webinars, Tutorials, and Bootcamps**

- Data management lessons learned
- Ingest best practices
- Data discovery and access

→ **ESGF Webinar series playlist at** <https://www.youtube.com/@esgf2432>

- **Hackathons and Workshops**

- Data standards
- Data node deployment and user compute resources
- Hold at large relevant conferences, e.g.,
AGU Fall Meeting, EGU, and AMS Annual Meeting

- Organize and host annual
ESGF Developer and User Conferences

→ **Ninth ESGF Developer and User Dual-Hybrid Conference held January 18–20, 2023**

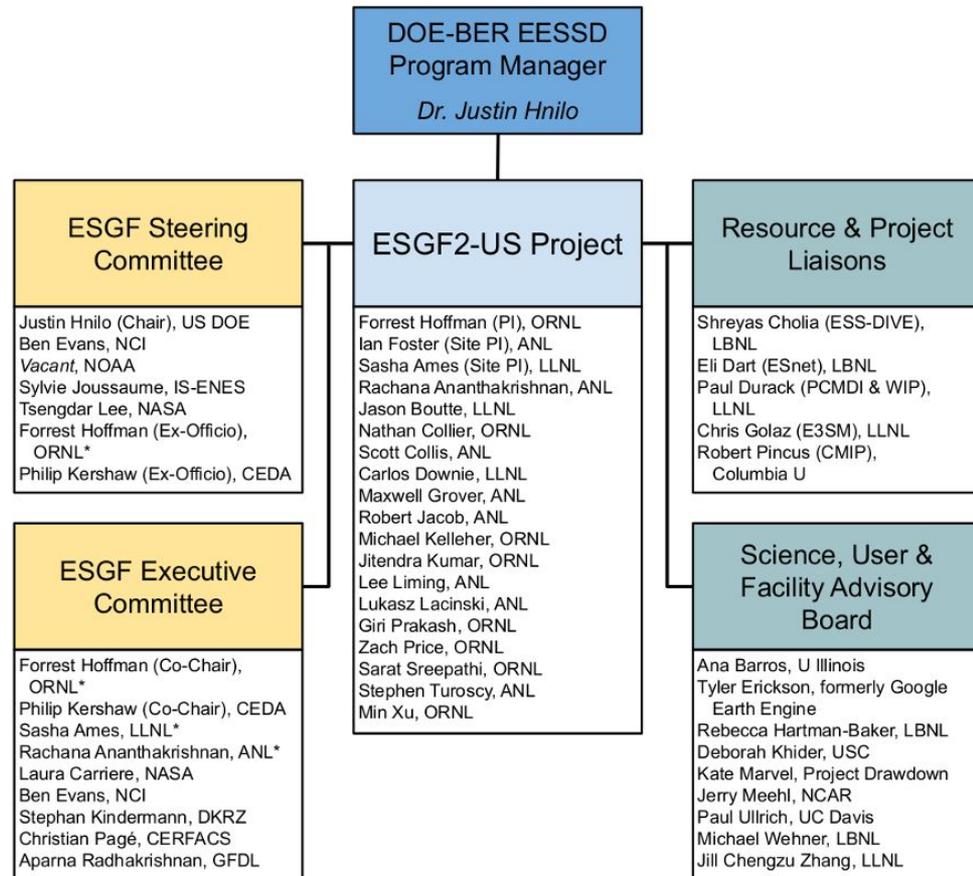


Ninth ESGF Developer and User Conference, held jointly between Oak Ridge National Laboratory (USA) and Toulouse (France), January 18–20, 2023



ESGF and ESGF2-US Project Governance

- ESGF is governed by an international **ESGF Executive Committee** that meets monthly
- The **XC** is directed by the **ESGF Steering Committee (SC)**, composed primarily of sponsoring agency representatives
- The ESGF2-US Project added
 - Resource & Project Liaisons** group to enhance communication with interdependent projects and critical resources
 - Science, User & Facility Advisory Board** to evaluate and prioritize project efforts with respect to community needs



*Indicates ESGF2-US team member

Updated April 11, 2023

The logo for ESGF US, featuring a stylized globe with the text 'ESGF' in large blue letters and 'US' in smaller blue letters below it.

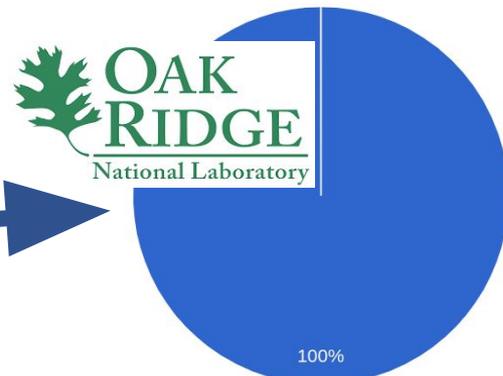
ESGF-US Failsafe Data Replication

- **In the US, LLNL operates the primary ESGF node**, which replicates much of the CMIP6 and related model output from around the globe
- Since the data at LLNL are contained only on spinning disk, we decided to replicate the **entire ~7.5 PB collection of data** to Argonne National Laboratory (ANL) and Oak Ridge National Laboratory (ORNL)
- **Solution: Use Globus to transfer all the data over ESnet**
- We used custom Globus scripting (*thanks to Lukasz Lacinski*), ESnet network monitoring and diagnostics (*thanks to Eli Dart*), DTN and GPFS optimized configurations (*thanks to Cameron Harr and others*), and debugging and problem-solving (*thanks to Sasha Ames, Lee Liming, and others*)



Data transferred to ALCF

Data transferred to OLCF



1.5 GB/s

4 to 6 GB/s



Replication to ALCF

ACTIVE, PAUSED and the latest SUCCEEDED transfers

7.5 PB transferred between mid-Feb and May 4
17,347,671 directories and 28,907,532 files

No	Datasets	From	Requested	Completed	Status	Directories	Files	Bytes Transferred	Faults	Rate
1	/cmip5_css01_data/cmip5/output1/NSF-DOE-NCAR/CESM1-CAM5	LLNL	2022-05-03 08:46:03	2022-05-04 11:37:43	SUCCEEDED	7208	13540	29913341340	16	309 kB/s
2	/cmip5_css02_data/cmip5/output1/NCC/NorESM1-M	LLNL	2022-05-02 09:52:03	2022-05-02 11:31:27	SUCCEEDED	4017	7548	5367692747060	0	900 MB/s
3	/cmip5_css02_data/cmip5/output1/NCAR/CCSM4	LLNL	2022-05-02 01:53:03	2022-05-03 00:50:23	SUCCEEDED	52571	48925	33455438769668	11	405 MB/s
4	/cmip5_css02_data/cmip5/output1/NASA-GISS/GISS-E2-R-CC	LLNL	2022-05-02 01:28:03	2022-05-02 01:52:31	SUCCEEDED	2098	9576	1087745609416	0	741 MB/s
5	/cmip5_css02_data/cmip5/output1/NASA-GISS/GISS-E2-R	LLNL	2022-05-02 00:42:03	2022-05-02 09:51:16	SUCCEEDED	30164	132059	24482369232188	5	743 MB/s

Replication to OLCF

ACTIVE, PAUSED and the latest SUCCEEDED transfers

No	Datasets	From	Requested	Completed	Status	Directories	Files	Bytes Transferred	Faults	Rate
1	/cmip5_css01_data/cmip5/output1/NSF-DOE-NCAR/CESM1-CAM5	LLNL	2022-05-03 08:47:18	2022-05-04 11:41:11	SUCCEEDED	7208	13540	271068730	16	2.80 kB/s
2	/cmip5_css02_data/cmip5/output1/NCAR/CCSM4	LLNL	2022-05-02 13:58:03	2022-05-03 03:14:27	SUCCEEDED	52571	48925	33455438769668	1	700 MB/s
3	/cmip5_css02_data/cmip5/output1/NCC/NorESM1-M	ALCF	2022-05-02 11:32:03	2022-05-02 12:15:48	SUCCEEDED	4017	7548	5367692747060	0	2.04 GB/s
4	/cmip5_css02_data/cmip5/output1/NASA-GISS/GISS-E2-R	ALCF	2022-05-02 09:52:03	2022-05-02 12:30:08	SUCCEEDED	30164	132059	24482369232188	3	2.58 GB/s
5	/cmip5_css02_data/cmip5/output1/NASA-GISS/GISS-E2-R-CC	ALCF	2022-05-02 05:34:04	2022-05-02 05:44:32	SUCCEEDED	2098	9576	1087745609416	0	1.73 GB/s



<https://dashboard.globus.org/esgf>

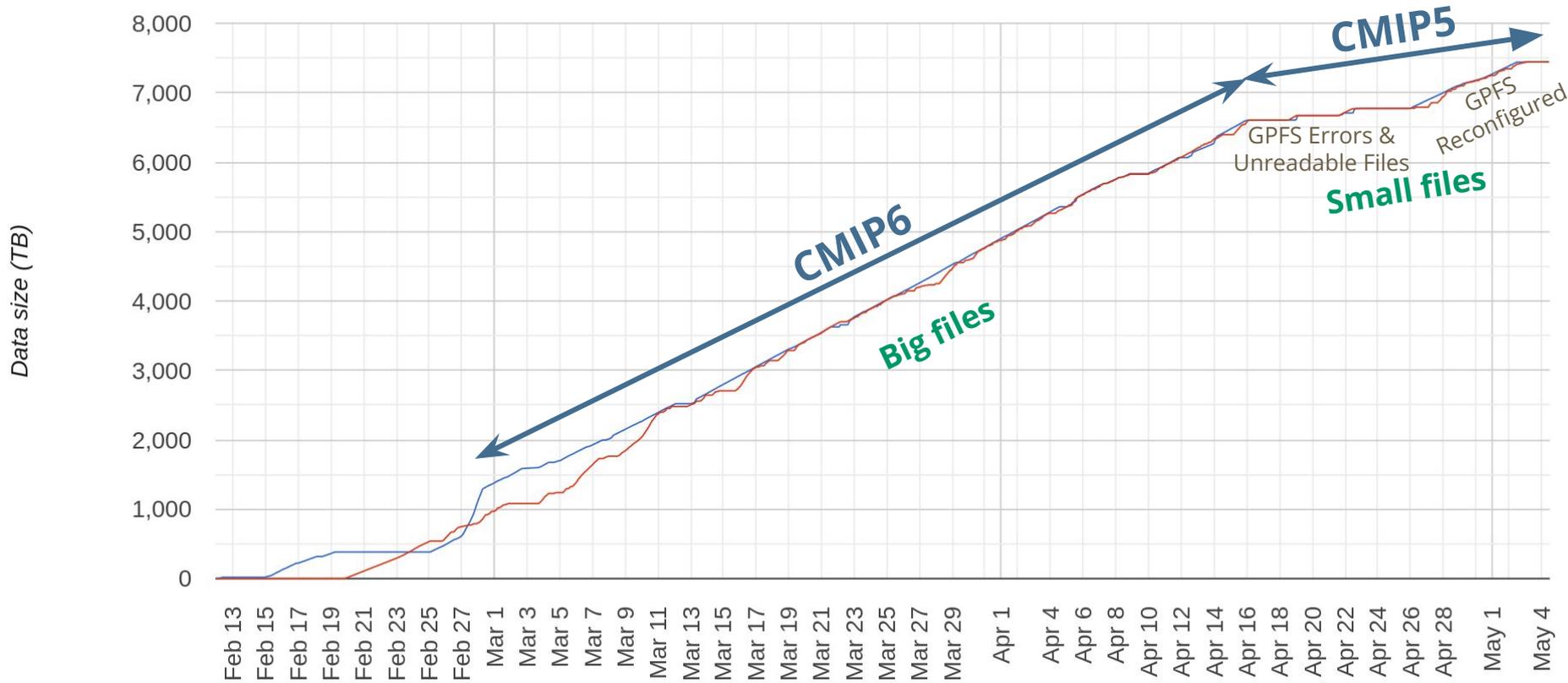
As of May 4, 2022



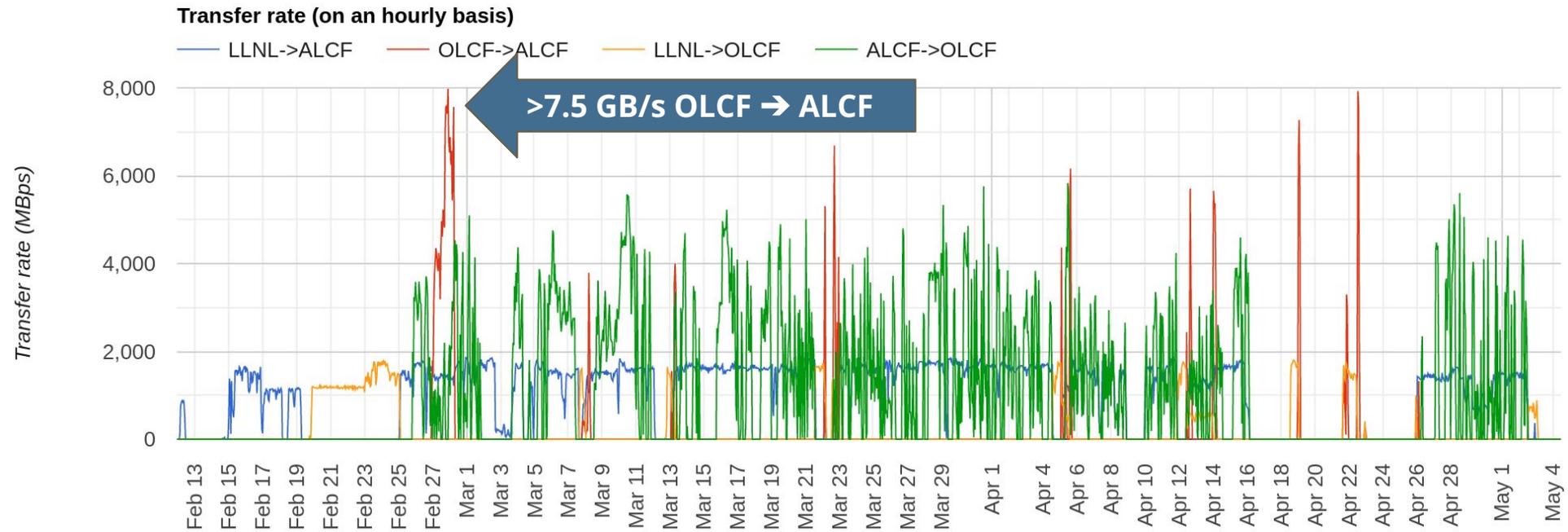
Cumulative Data Transferred Over Time

Progress of transfers

— to ALCF — to OLCF



ESGF² US Transfer Rates Over Time



ESGF2 ARM Data Center

ARM Research Facility | arm.gov/data/

ARM | U.S. DEPARTMENT OF ENERGY | Office of Science

DATA | CAPABILITIES | RESEARCH | NEWS & EVENTS | ABOUT

Data Discovery | Data Sources | Work with ARM Data | Data Quality Program

DATA

Atmospheric Radiation Measurement (ARM) user facility data are collected through routine operations and scientific field experiments.

ARM focuses on obtaining continuous measurements—supplemented by field campaigns—and providing data products that promote the advancement of climate models. Serving users worldwide, the ARM Data Center collects and archives approximately 20 terabytes of data per month. Datastreams are generally available for download within 48 hours.

ABOUT ARM DATA

ARM data include routine data products, value-added products (VAPs), field campaign data, complementary external data products from collaborating programs, and data contributed by ARM principal investigators for use by the scientific community. Data quality reports, graphical displays of data availability/quality, and data plots are also available from the ARM Data Center.

As a U.S. Department of Energy national scientific user facility, ARM welcomes users from all institutions and nations. A free ARM user account is needed to access ARM data, a requirement of the Office of Management and Budget.

CREATE ACCOUNT | UPDATE ACCOUNT

FINDING AND USING ARM DATA

All data obtained through ARM are monitored for quality and made available **free of charge** through the ARM Data Center via [Data Discovery](#).

GET ARM DATA | Search ARM Data... | Q

FEATURED DATA

Updated Stereo Camera Cloudiness Map Product Released for Evaluation
20 March 2023

Raman Lidar Vertical Profiles Product Expands to Oriskany Point, Alaska

ARM Data Discovery | adc.arm.gov/discovery/#/

ARM | HOME | DATA | SUPPORT | CART

Click to interact

Data Search
Search by category, measurement, datastream + more.

Guided Search
For new users, search through guided questions.

Location Search Beta
Use the map to search by site or facility.

Data Search

Select one of the categories below or type in the search text box to get started.

Enter a category, measurement, datastream, site, source or keyword to begin your search.

Atmospheric Radiation Measurement (ARM) Data Center – <https://www.arm.gov/data/>

ESGF US 2 ESS Deep Insight for Earth Science Data (DIVE)

The screenshot shows the homepage of the ESS-DIVE website. At the top left is the ESGF US logo. The main header features the ESS-DIVE logo and the tagline "Deep Insight for Earth Science Data". Below this are navigation links for DATA, ABOUT, GET STARTED, and LEARN MORE. A large central image shows a pair of hands holding a globe of the Earth. To the left of the globe are two buttons: "Search Data" and "Submit Data". Below the main image, there is a "LATEST" section with a thumbnail for a "February Webinar on ESS-DIVE..."

The screenshot shows the search results page for the ESS-DIVE Data Archive. The URL is data.ess-dive.lbl.gov/data. The page has a navigation bar with links for DATA, PORTALS, PROJECTS, GET STARTED, ABOUT, and a prominent "SUBMIT DATA" button. A "Sign in with Orcid" button is also present. The search interface includes a search bar, a "Sort by" dropdown set to "Most recent", and a pagination control showing "1 2 3 ... 30 Next". A "Filter by" sidebar lists categories: Project, Identifier, Region description, Creator, Year, and Access. The main content area displays search results for a dataset titled "Air temperature and relative humidity raw data from June 2017-Oct 2018 at the B34 tower in Manaus, Brazil". The results list authors (de Jesus Sampaio Filho I, Candido L, Araujo A, Gimenez B, Higuchi N, Chambers J) and provide a DOI: 10.15486/NGT/1602481. A second result is also visible, titled "WHONDRS 48 Hour Diel Cycling Study at the Nisqually River, WA". On the right side of the page is a world map with a grid overlay, showing the location of the data points. The map includes a legend for "Satellite" and "Terrain" and a scale bar for 1000 km.

Environmental Systems Science-Deep Insight for Earth Science Data (ESS-DIVE) – <https://ess-dive.lbl.gov/>

ESGF US 2 A Unified Data Framework for DOE BER (RFI)

DOE's Biological and Environmental Research Advisory Committee (BERAC) is collecting information on ways to integrate existing data activities based on charge from Undersecretary of SC

Federal Register
PDF
<https://www.govinfo.gov/content/pkg/FR-2023-04-17/pdf/2023-08029.pdf>
Webpage
<https://www.federalregister.gov/documents/2023/04/17/2023-08029/a-unified-data-framework-for-doe-biological-and-environmental-research>

Federal Register / Vol. 88, No. 73 / Monday, April 17, 2023 / Notices **23415**

messaging and training are not resonating with Marines. In order to examine this trend within the Marine Corps and gather relevant baseline data for the larger DSPO initiative, GNA, in conjunction with the Marine Corps Marine and Family Programs Division (MFP), propose the information collection, "Marine Corps Safety Needs Assessment" survey. This voluntary survey examines current LMS program awareness, preferences for safety devices and locations, and the place of safety in Marine Corps culture. This survey will assist MFP in identifying, from the perspective of Marines, the reach of current LMS efforts and the acceptability of potential LMS activities. The results of the survey will be used by MFP and DSPO to better understand which LMS activities and messages resonate with Marines, as well as serve as a baseline data for future LMS activity effectiveness evaluations in accordance with the standards of practice framework prescribed by DoD Instruction 6490.16.

Affected Public: Individuals or households.
Annual Burden Hours: 2,262.
Number of Respondents: 9,048.
Responses per Respondent: 1.
Annual Responses: 9,048.
Average Burden per Response: 15 minutes.
Frequency: Once.
Dated: April 12, 2023.

Kayyome T. Marston,
Alternate OSD Federal Register Liaison Officer, Department of Defense.
[FR Doc. 2023-08050 Filed 4-14-23; 8:45 am]
BILLING CODE 5001-06-P

DEPARTMENT OF ENERGY
A Unified Data Framework for DOE Biological and Environmental Research

AGENCY: Office of Biological and Environmental Research (BER), Office of Science, Department of Energy (DOE).
ACTION: Request for information.

SUMMARY: The Biological and Environmental Research (BER) Program, as DOE's coordinating office for research on biological systems, bioenergy, environmental science, and Earth system science, is seeking input on the need and the structure of a unified data framework that links or integrates existing data activities within BER. Information produced in response to this request may be used by the BER Advisory Committee (BERAC) to help inform and recommend to BER a strategy for next-generation data management and analysis within a unified framework.

Information is specifically requested on how a more unified data infrastructure may better facilitate current or future science questions, and what components or technologies are needed to develop a more unified data infrastructure. Answers or information related, but not limited to, the following questions are specifically requested:

1. Do you conduct research in one of the BER research areas (i.e., Atmospheric Science; Earth and Environmental System Modeling; Environmental Science; Bioenergy and Bioproducts; or Plant and Microbial Genomics) and, if so, which area(s)? Please limit additional detail on your area(s) of research interest to a brief paragraph.
2. What new or existing research areas might benefit from improvements in data availability or access across research areas, potentially enabling scientific breakthroughs—and why?
3. What data improvements, including those of accessibility and integration, could facilitate new or existing research or scientific breakthroughs?
 - a. Are there current data sets that should be linked or integrated into existing data infrastructure to facilitate existing or new research? If so, which data sets should be so linked or integrated and why?
 - b. Are there current barriers to accessing or integrating data from (a) different DOE sources (e.g., ARM, JGI, ESS-DIVE, MSD-LIVE) or from (b) different sources separately maintained by DOE and another Federal agency? If so, what are those barriers and how might they be addressed to allow for improved data access and integration?
 - c. What data infrastructure improvements would best support model-experiment feedbacks; facilitate data synthesis and analysis for multidisciplinary research; and enable application of advanced statistical techniques, including artificial intelligence and machine learning? Please include a brief explanation as to how each identified improvement would support each of these listed tasks.
 - d. What current barriers need to be addressed in developing a unified infrastructure to promote greater use by a more diverse community of users, with a focus on improving diversity, equity, and inclusion within data usage and application?

While the questions provided above can help guide thinking on this topic, any input is welcome that may assist BERAC in developing a next-generation data infrastructure in support of BER mission science. The information provided through this request will assist

FEDERAL REGISTER
The Daily Journal of the United States Government

A Unified Data Framework for DOE Biological and Environmental Research

A Notice by the Energy Department on 04/17/2023

This document has a comment period that ends in 176 days. (10/31/2023)

SUBMIT A FORMAL COMMENT

1 comments received. [View posted comments](#)

PUBLISHED DOCUMENT

AGENCY: Office of Biological and Environmental Research (BER), Office of Science, Department of Energy (DOE).

ACTION: Request for information.

SUMMARY: The Biological and Environmental Research (BER) Program, as DOE's coordinating office for research on biological systems, bioenergy, environmental science, and Earth system science, is seeking input on the need and the structure of a unified data framework that links or integrates existing data activities within BER. Information produced in response to this request may be used by the BER Advisory Committee (BERAC) to help inform and recommend to BER a strategy for next-generation data management and analysis within a unified framework.

DATES: Written comments and information are requested on or before October 31, 2023.

DOCUMENT DETAILS

Printed version: PDF
Publication Date: 04/17/2023
Agency: Department of Energy
Dates: Written comments and information are requested on or before October 31, 2023.
Comments Close: 10/31/2023
Document Type: Notice
Document Citation: 88 FR 23415
Page: 23415-23416 (2 pages)
Document Number: 2023-08029

DOCUMENT DETAILS



ESGF2 Summary and Integration Activities

- The next generation **Earth System Grid Federation (ESGF2-US)**
 - Will be designed for an **order of magnitude increase in data sizes**
 - Will be highly **available, scalable, and fast**
 - Will have improved **data discovery and sharing tools**
 - Will offer **server-side computing** for derived data
 - Will offer **user computing capabilities** (e.g., JupyterHub/JupyterLab) near the data
 - Will be developed collaboratively with **international Federation partners**
- All new **ESGF development is being performed collaboratively** with Federation partners all over the world
- ESGF2-US is integrated with **DOE Earth & Environ. Systems Modeling projects** and with international **WCRP CMIP activities**, including serving on multiple Task Teams for CMIP7
- ESGF2-US aims to add **new data projects** to support large-scale AI/ML data, multi-agency model intercomparisons, and model benchmarking
- **User computing** approaches initiated in the commercial cloud and deployed through on-premise cloud infrastructure will likely facilitate more rapid research and discovery