The Next Generation Earth System Grid Federation

Forrest M. Hoffman (Oak Ridge National Laboratory)

National Center for Atmospheric Research (NCAR)
Computational and Information Systems Laboratory (CISL) Seminar

April 11, 2023

U.S. DEPARTMENT OF ENERGY | Office of Science

Forrest M. Hoffman, Computational Earth System Scientist

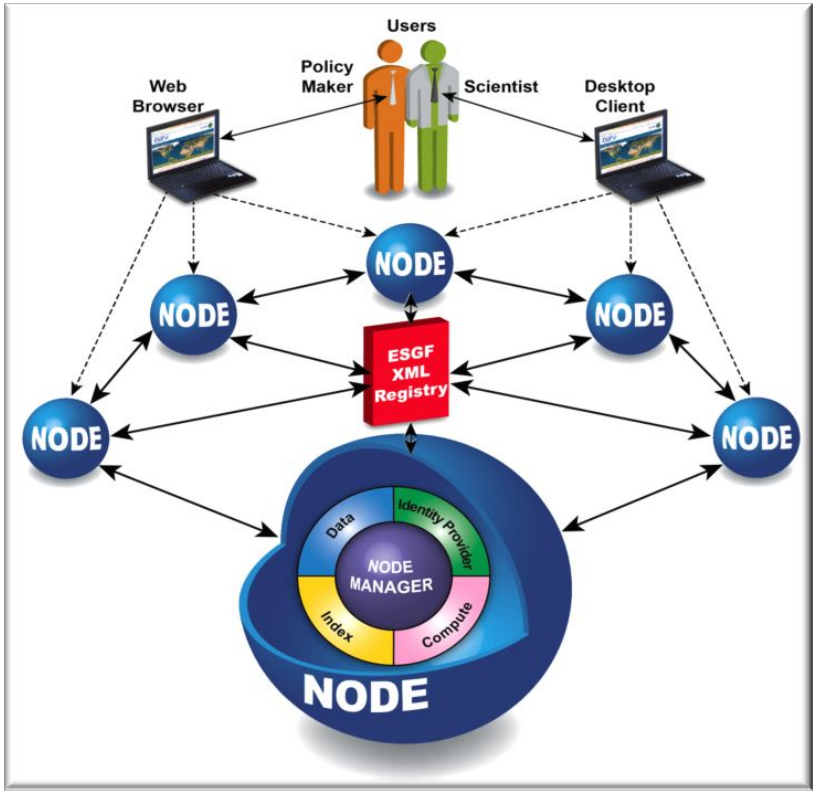
- Group Leader for the ORNL Computational Earth Sciences Group
- 34 years at ORNL in Environmental Sciences Division, then Computer Science and Mathematics Division, and now Computational Sciences and Engineering Division
- Develop and apply Earth system models to study global biogeochemical cycles, 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

OAK RIDGE National Laboratory

THE UNIVERSITY OF TENNESSEE KNOXVILLE

What is the Earth System Grid Federation?

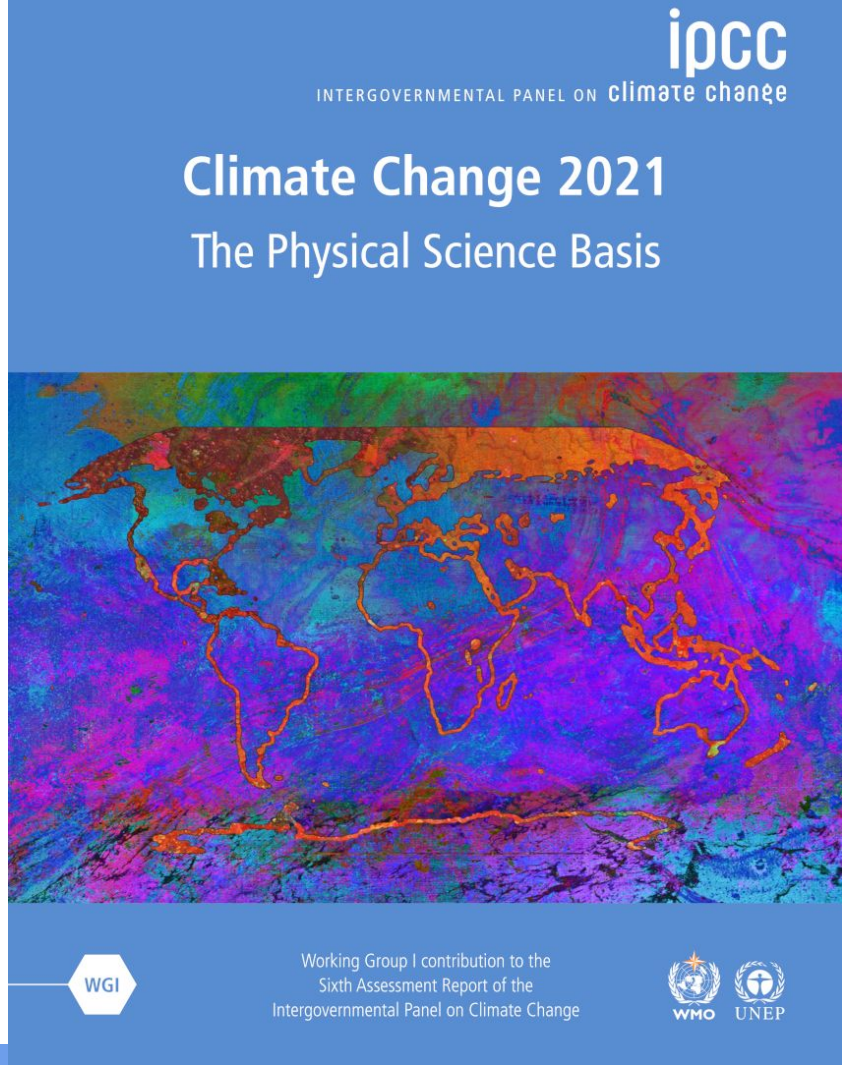
- The **Earth System Grid Federation (ESGF)** is an *international consortium* and a *globally distributed peer-to-peer network of data servers* using a common set of protocols and interfaces to archive and distribute climate and Earth system model output and related input, observational, and reanalysis data
- These 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

ESGF2 US IPCC AR6

- 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!**



ESGF Holdings are Large and Growing

- CMIP5 totals >5 PB (including replicas)
- CMIP6 totals >25 PB (including replicas)
- CMIP7 is expected to have more high resolution output & ensembles, totaling ~100 PB
- ESGF2-US will expand Federation holdings by adding other Earth science data projects for AI/ML, large ensembles, etc.

CMIP6	13,837,445 total datasets 25,463.24 TB	6,874,934 distinct datasets 14,437.12 TB	6,962,511 replica datasets 11,026.13 TB
CORDEX	185,894 total datasets 1,459.51 TB	185,622 distinct datasets 1,458.95 TB	272 replica datasets 0.56 TB
CMIP5	201,130 total datasets 5,293.61 TB	52,163 distinct datasets 1,525.07 TB	148,967 replica datasets 3,768.55 TB
INPUT4MIPS	11,504 total datasets 19.91 TB	5,660 distinct datasets 9.97 TB	5,844 replica datasets 9.95 TB
OBS4MIPS	210 total datasets 0.2 TB	210 distinct datasets 0.2 TB	0 replica datasets 0 TB

As of April 11, 2023

ESGF2 US A New Consortium Project in the USA

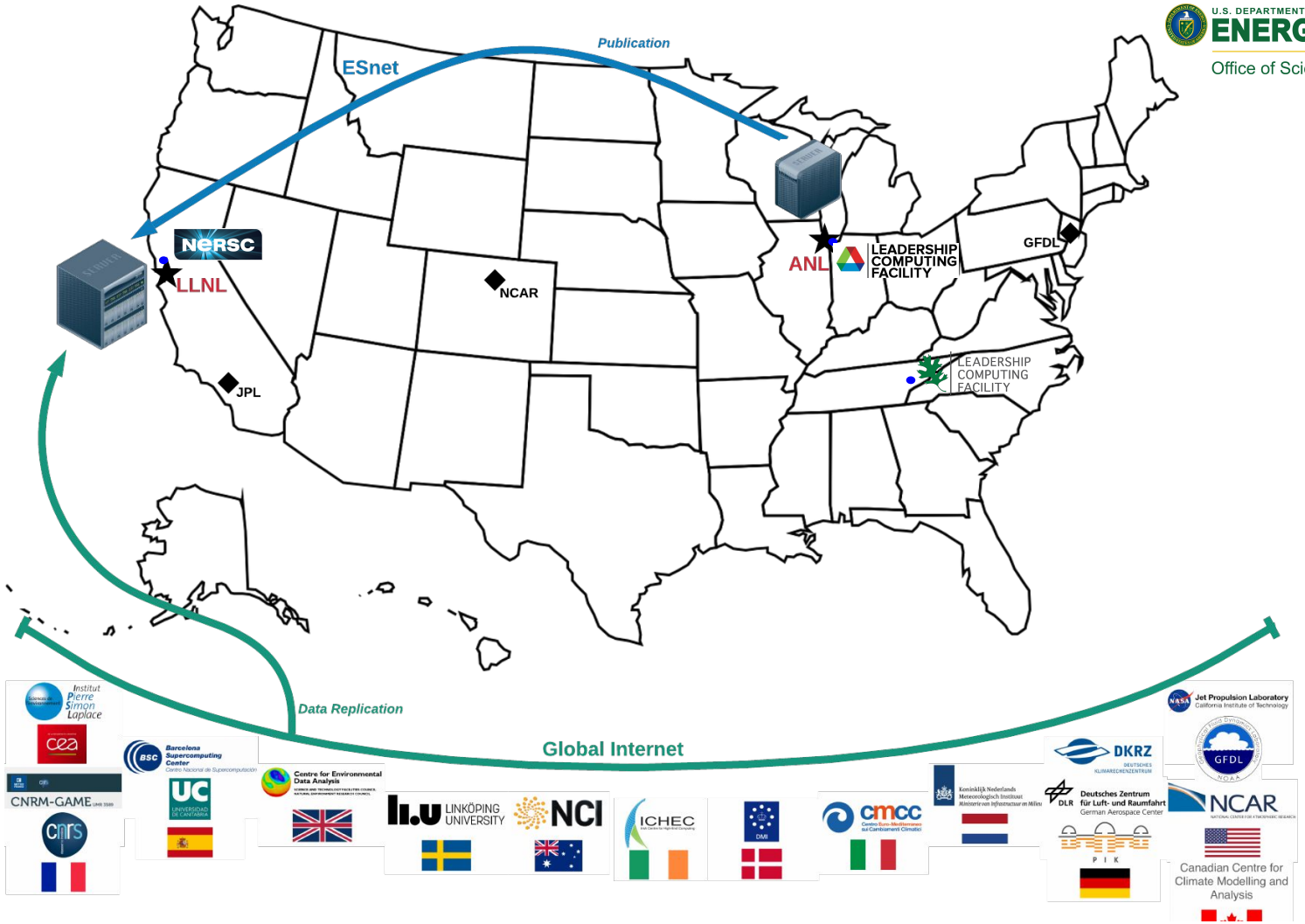
- New team from **Oak Ridge National Laboratory**, **Argonne National Laboratory**, and **Lawrence Livermore National Laboratory** proposed to modernize the data backplane based on the Globus platform
- ESGF2-US proposal was reviewed by panel of 8 scientists in August 2021, and was **selected for funding** 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

ESGF2 US

U.S. DEPARTMENT OF ENERGY
Office of Science

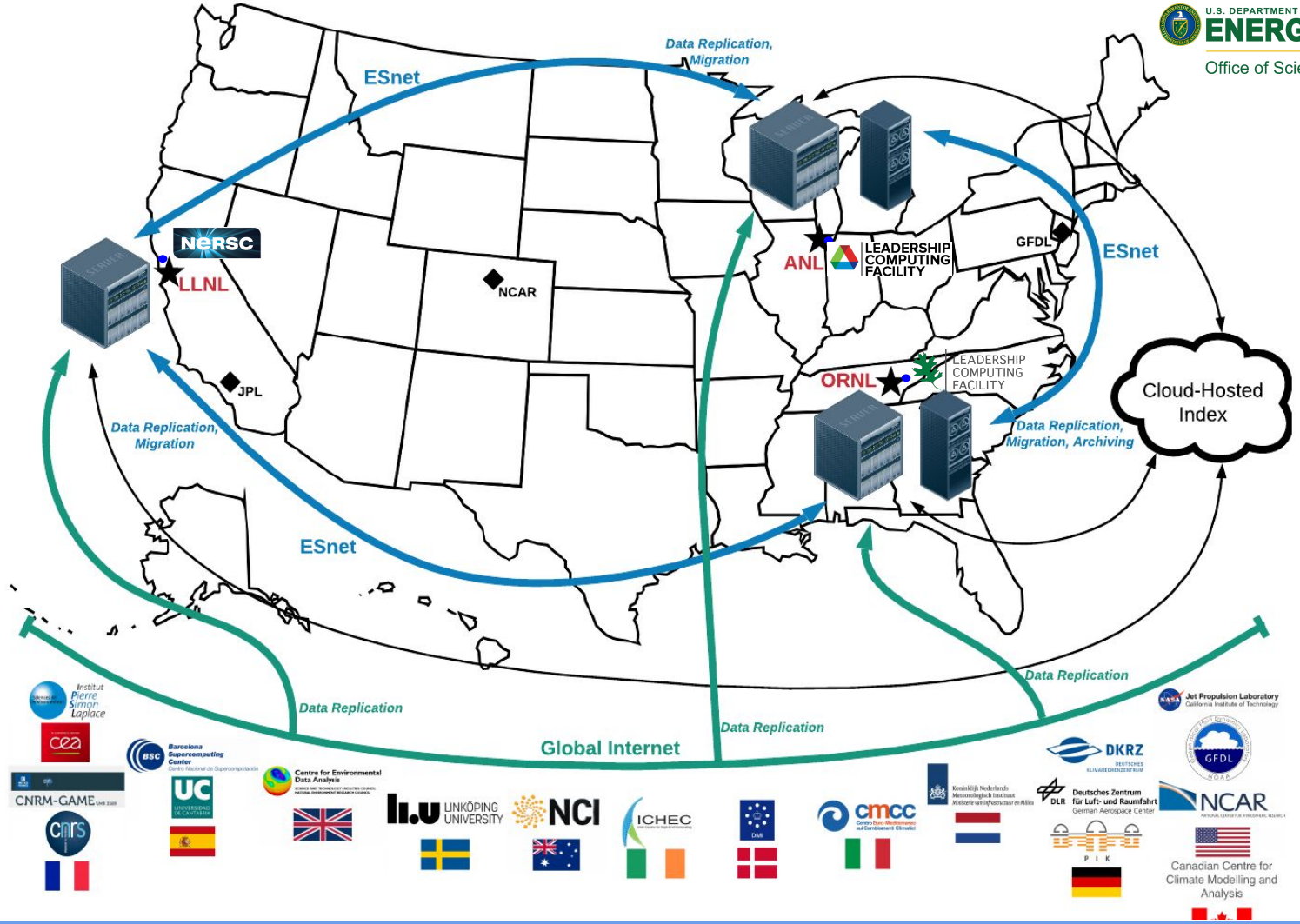
DOE's Current Earth System Grid Federation

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



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

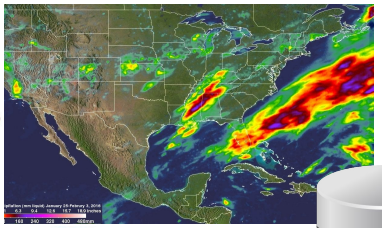


Design and implementation principles

- **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

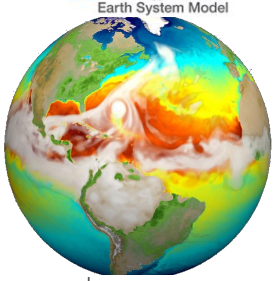
ESGF2 US Enabling a new level of research productivity

*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.*



ESGF2

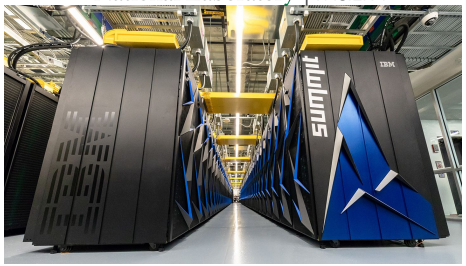
E3SM Energy Exascale Earth System Model

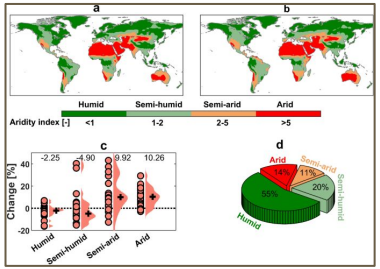


GitHub

jupyter

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY





Flood risk increases with water availability

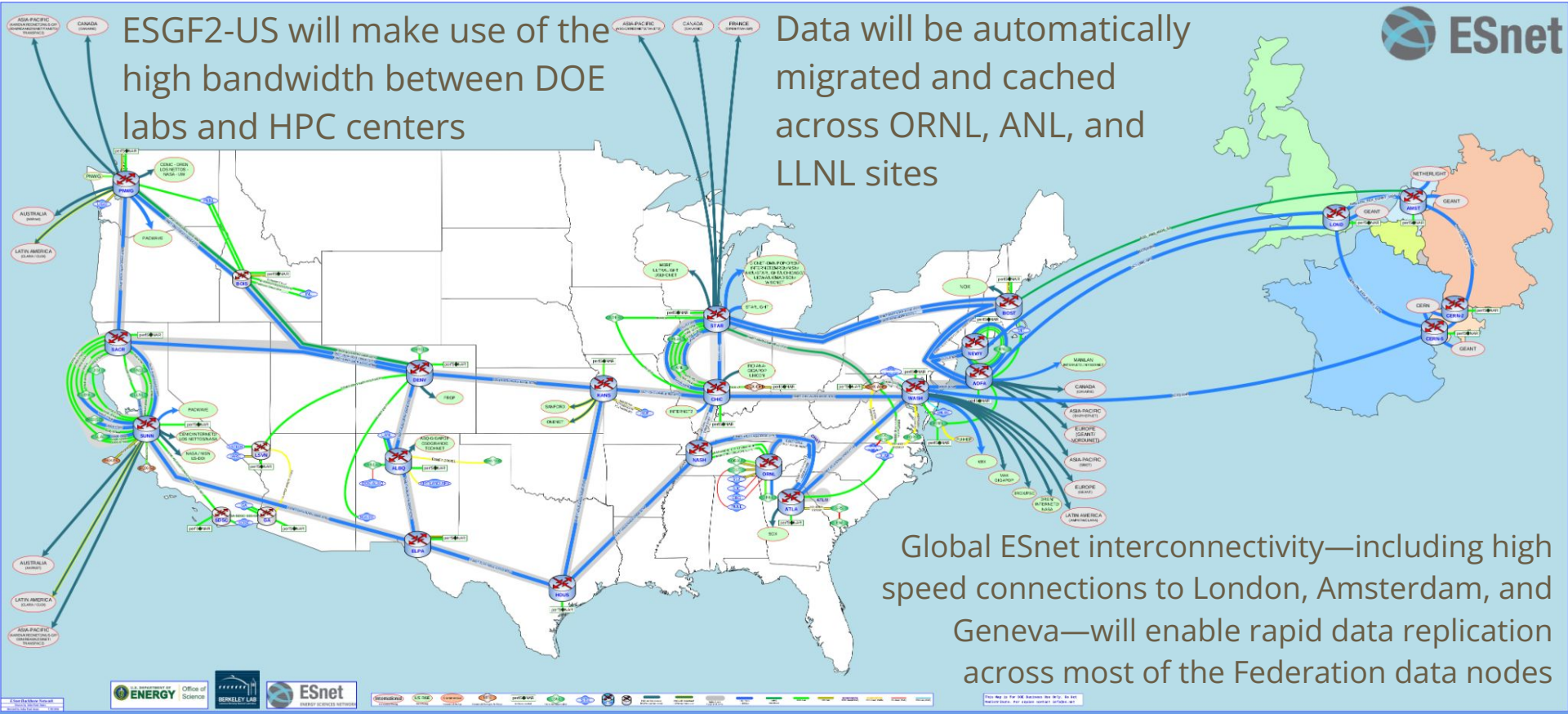
ESGF2 US ESnet Global Connectivity

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

ESnet
ENERGY SCIENCES NETWORK

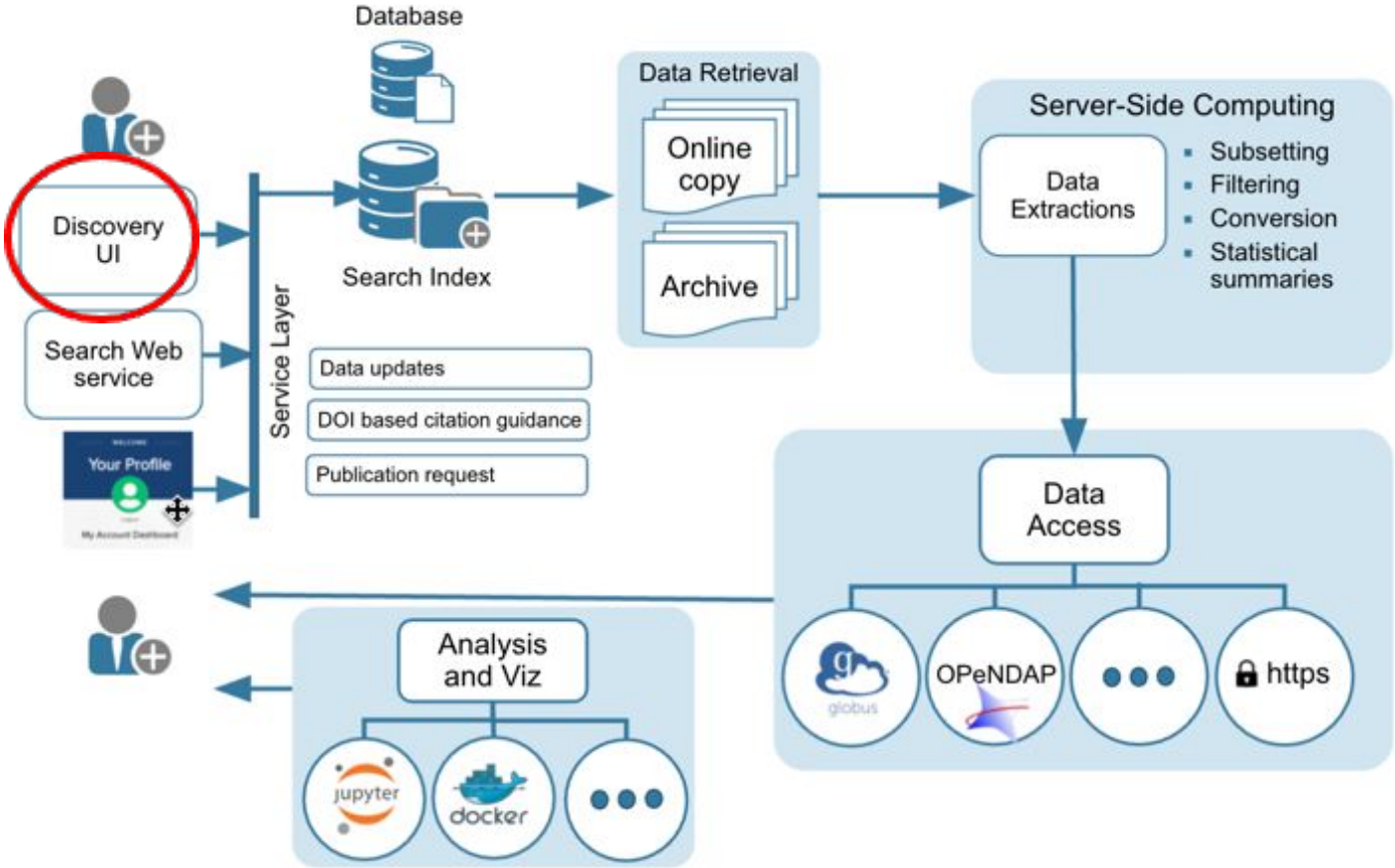
ESGF2-US will make use of the high bandwidth between DOE labs and HPC centers

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 most of the Federation data nodes

ESGF 2 US Data Discovery Platform: Architecture



ESGF2 US 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

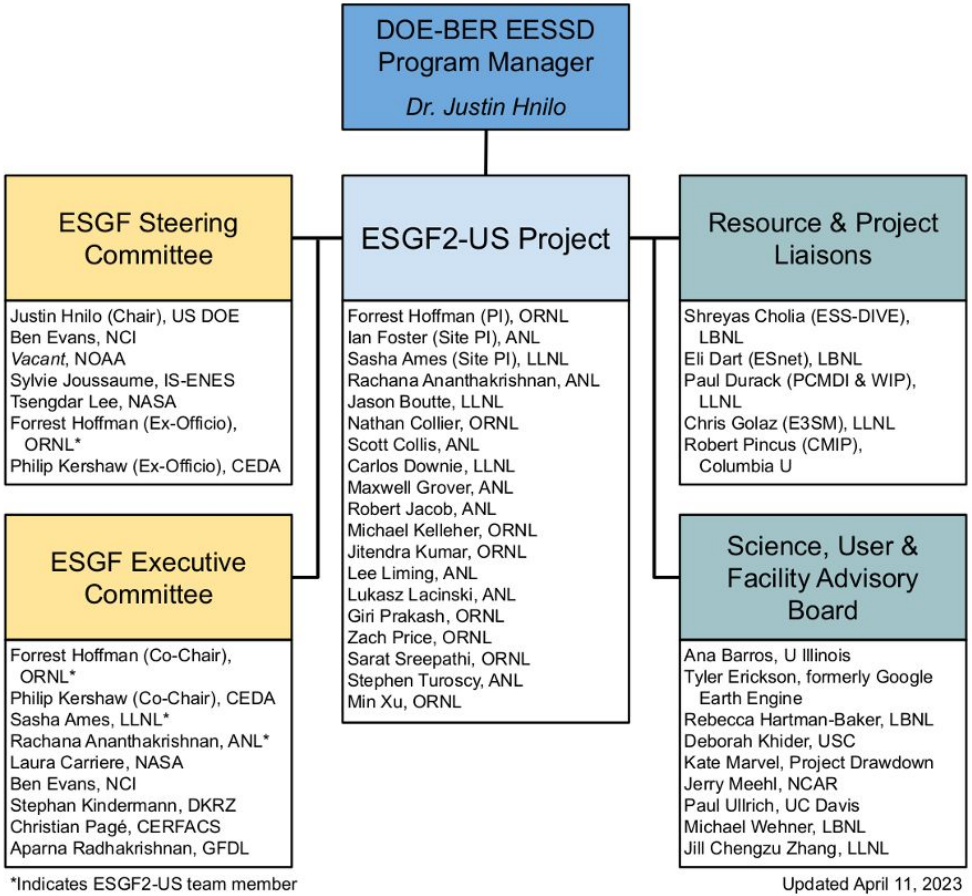
- Organize and host annual **ESGF Developer and User Conferences**

➜ **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



ESGF-US 2 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*)

ESGF2 US

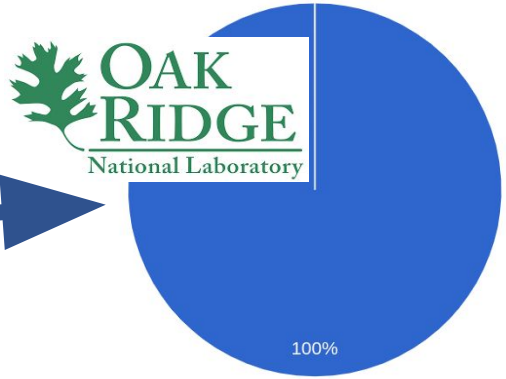
Data transferred to ALCF

1.5 GB/s

Lawrence Livermore National Laboratory

4 to 6 GB/s

Data transferred to OLCF



7.5 PB transferred between mid-Feb and May 4

17,347,671 directories and 28,907,532 files

Replication to ALCF

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 08:46:03	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

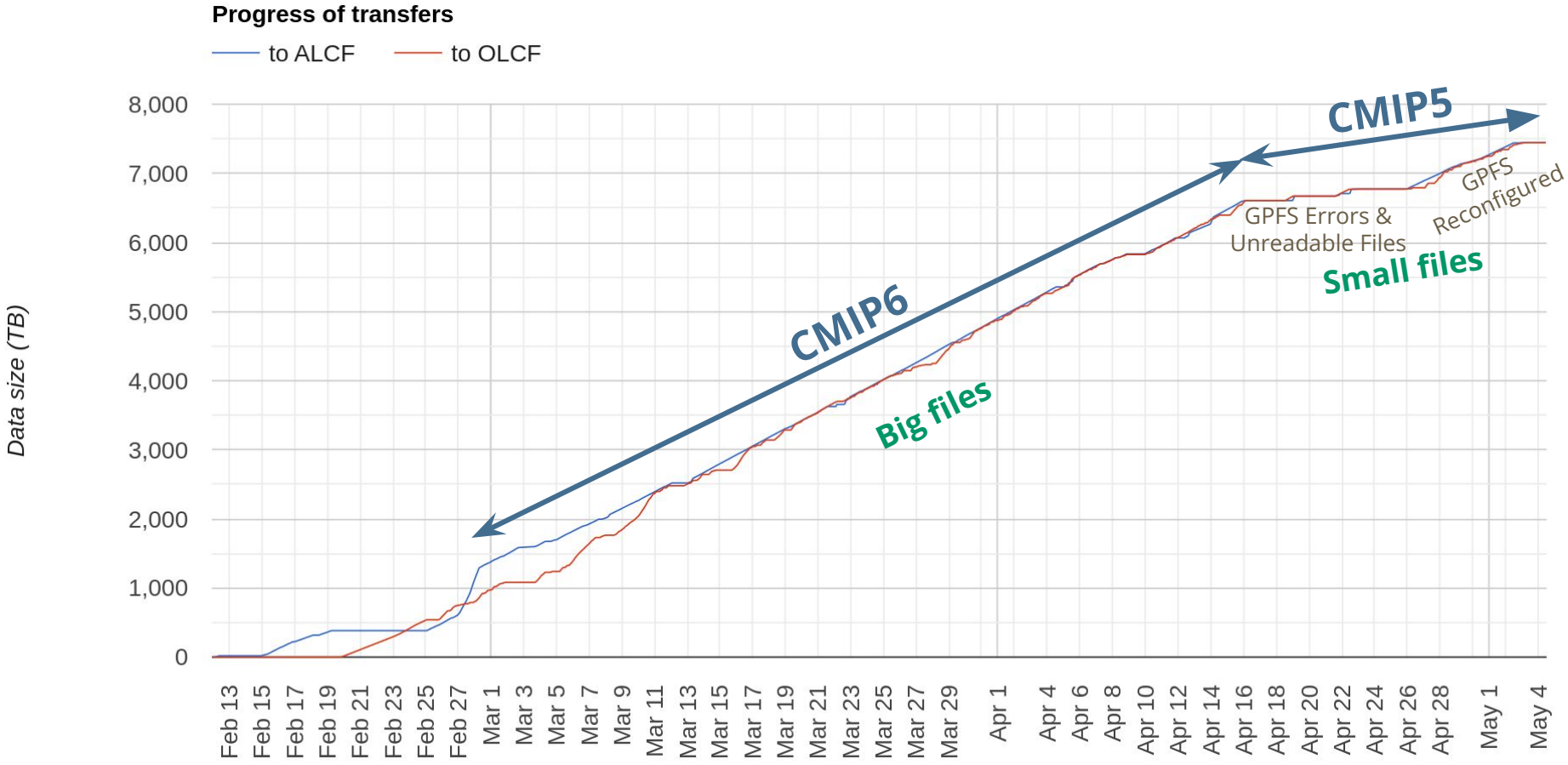
ESnet ENERGY SCIENCES NETWORK

globus

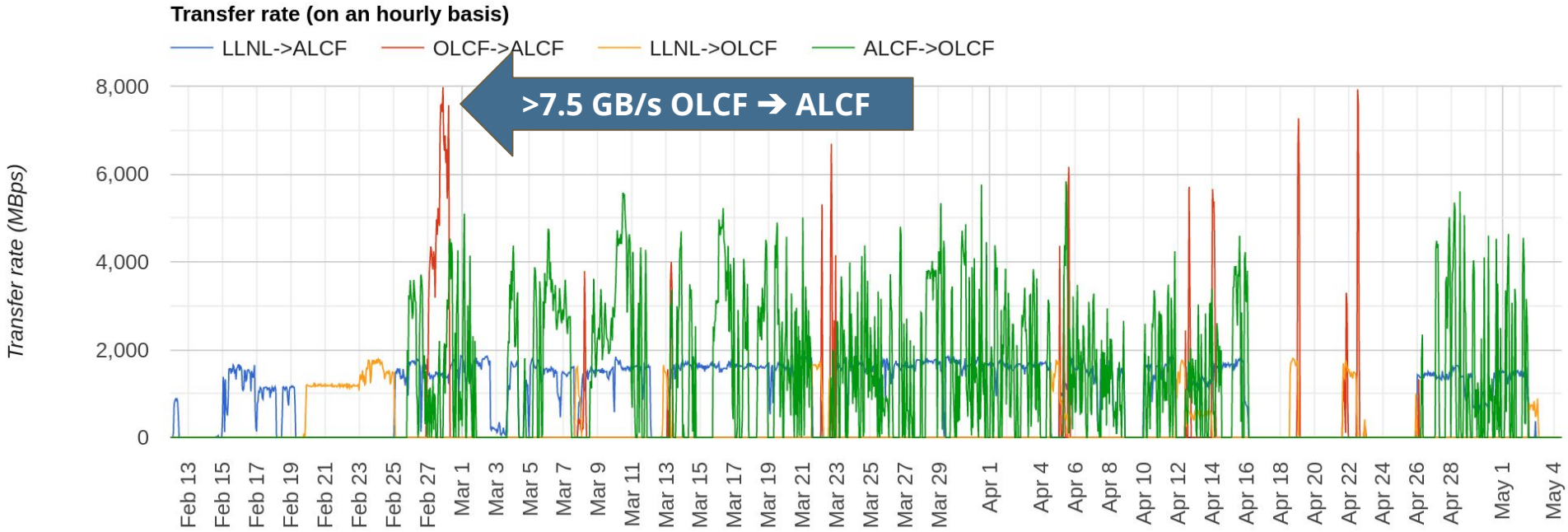
https://dashboard.globus.org/esgf

As of May 4, 2022

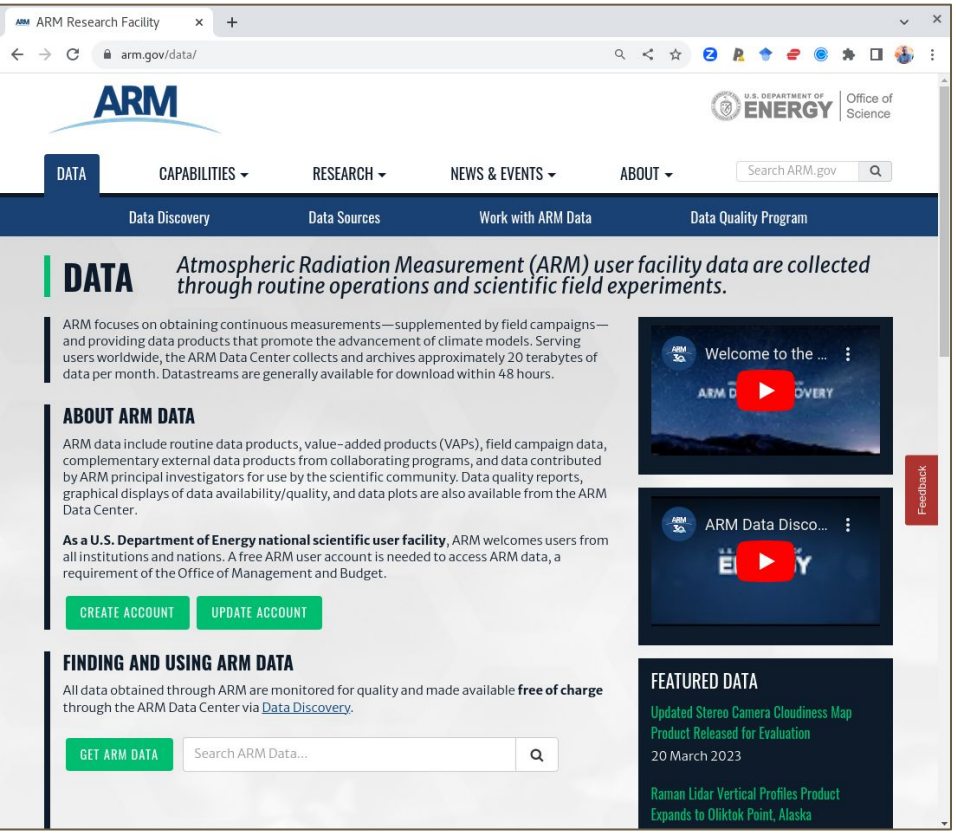
ESGF2 US Cumulative Data Transferred Over Time

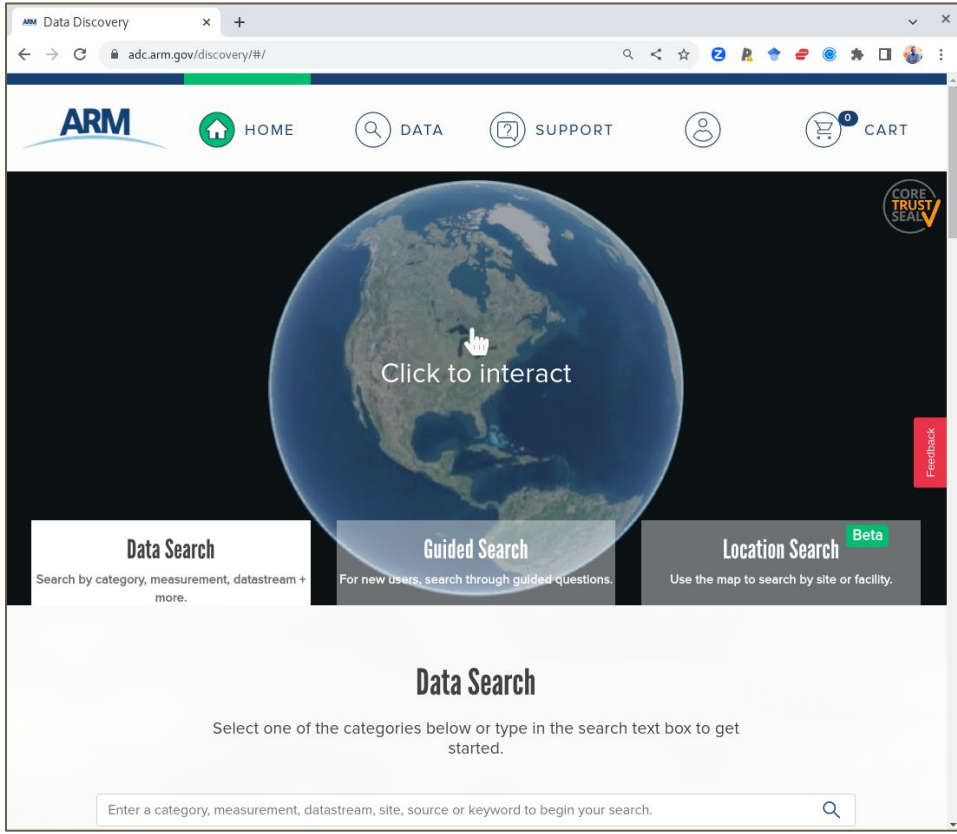


ESGF2 US Transfer Rates Over Time



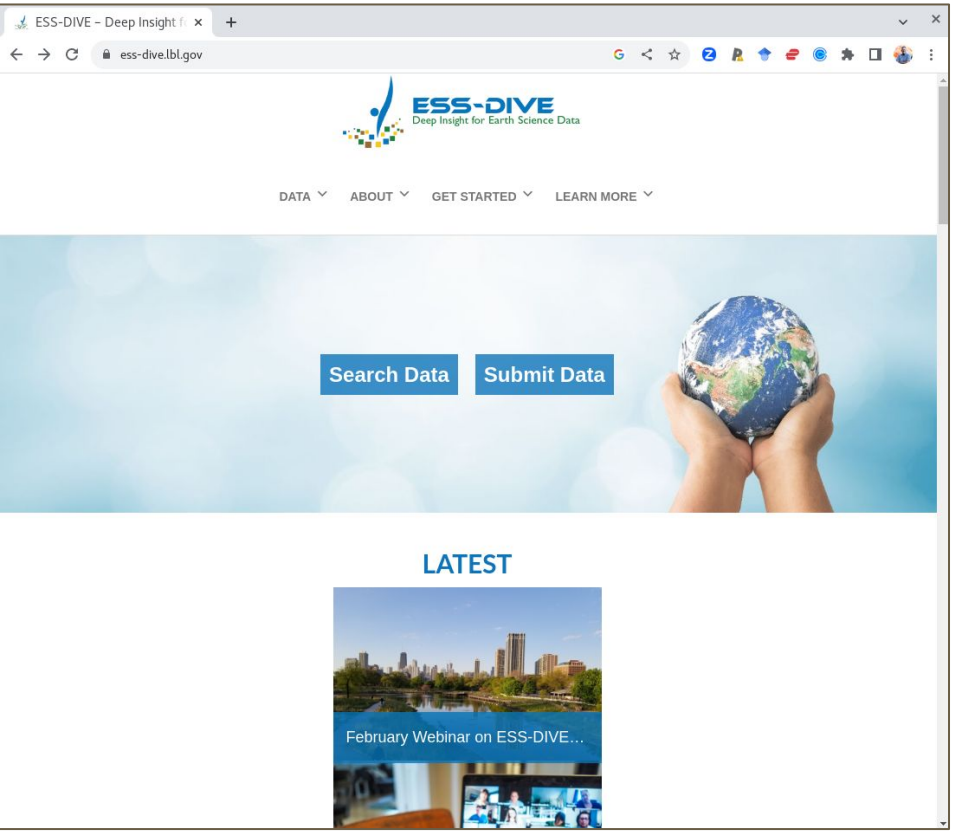
ESGF 2 US ARM Data Center

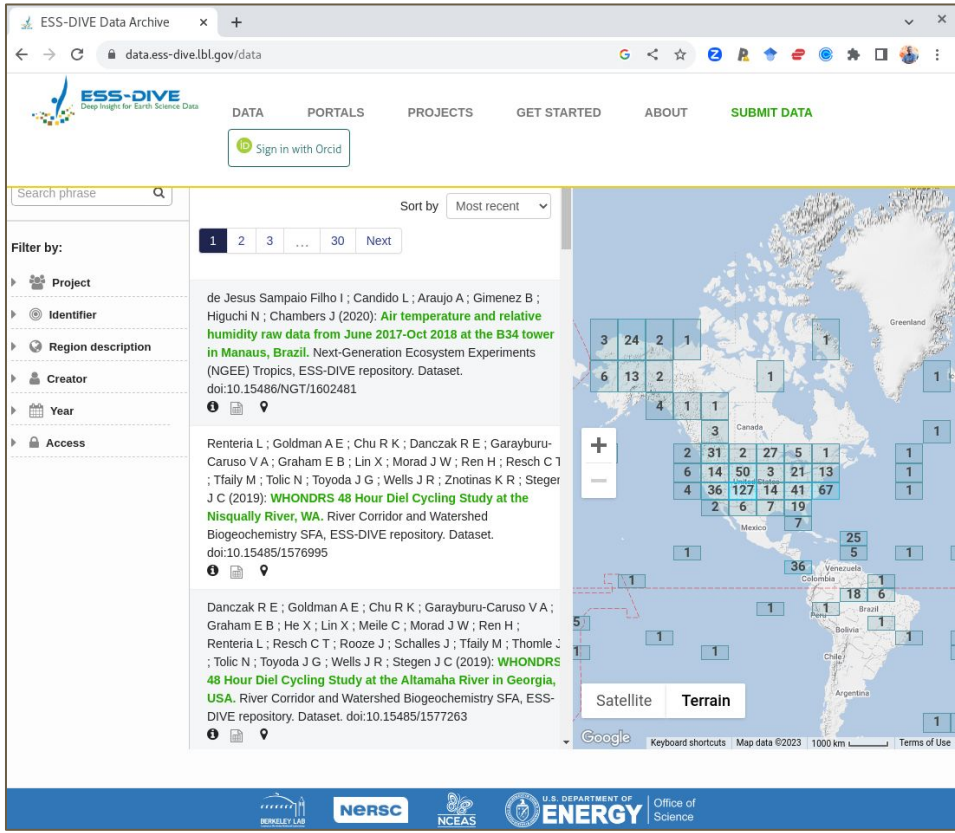




Atmospheric Radiation Measurement (ARM) Data Center – https://www.arm.gov/data/

ESGF2 US ESS Deep Insight for Earth Science Data (DIVE)





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

ESGF2 US Summary

- 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 automatically migrate data as needed
  - 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
- All new **ESGF development is being performed collaboratively** with Federation partners all over the world
- 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 research