EARTH SYSTEM GRID FEDERATION

AGU FALL MEETING
SCIENCE LEADS THE FUTURE
CHICAGO, IL & ONLINE EVERYWHERE
12-16 DECEMBER 2022

IN13A-01 - ESGF2-US: Building the Next Generation Earth System Grid Federation

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

2022 AGU Fall Meeting, Chicago, Illinois

U.S. DEPARTMENT OF ENERGY | Office of Science

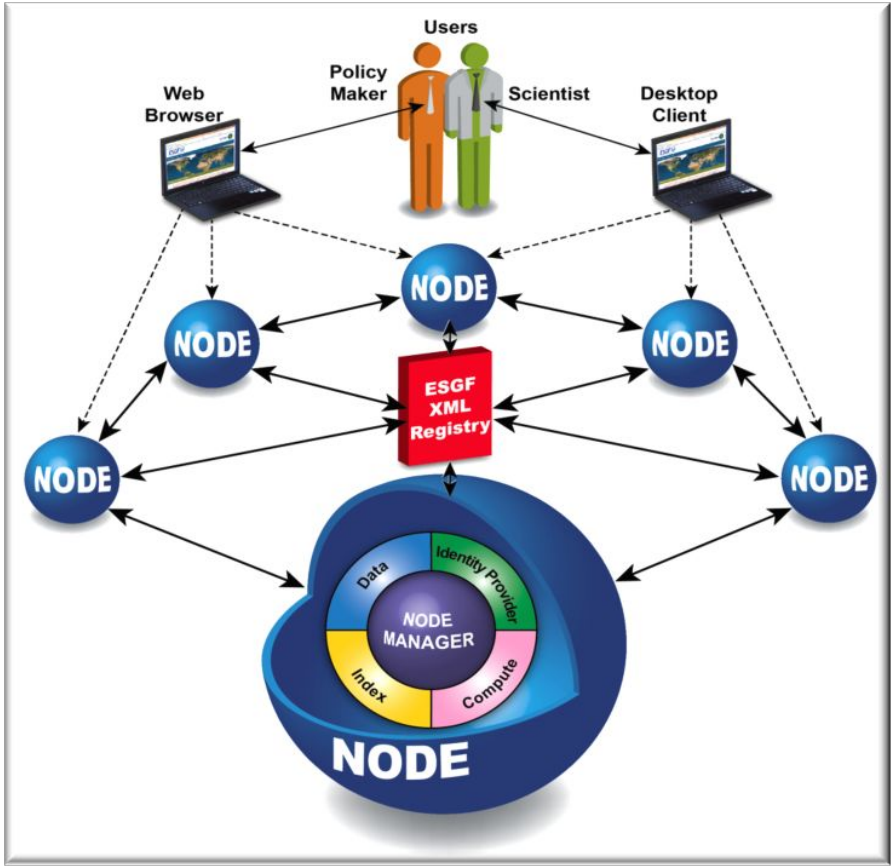
OAK RIDGE National Laboratory

Argonne NATIONAL LABORATORY

Lawrence Livermore National Laboratory

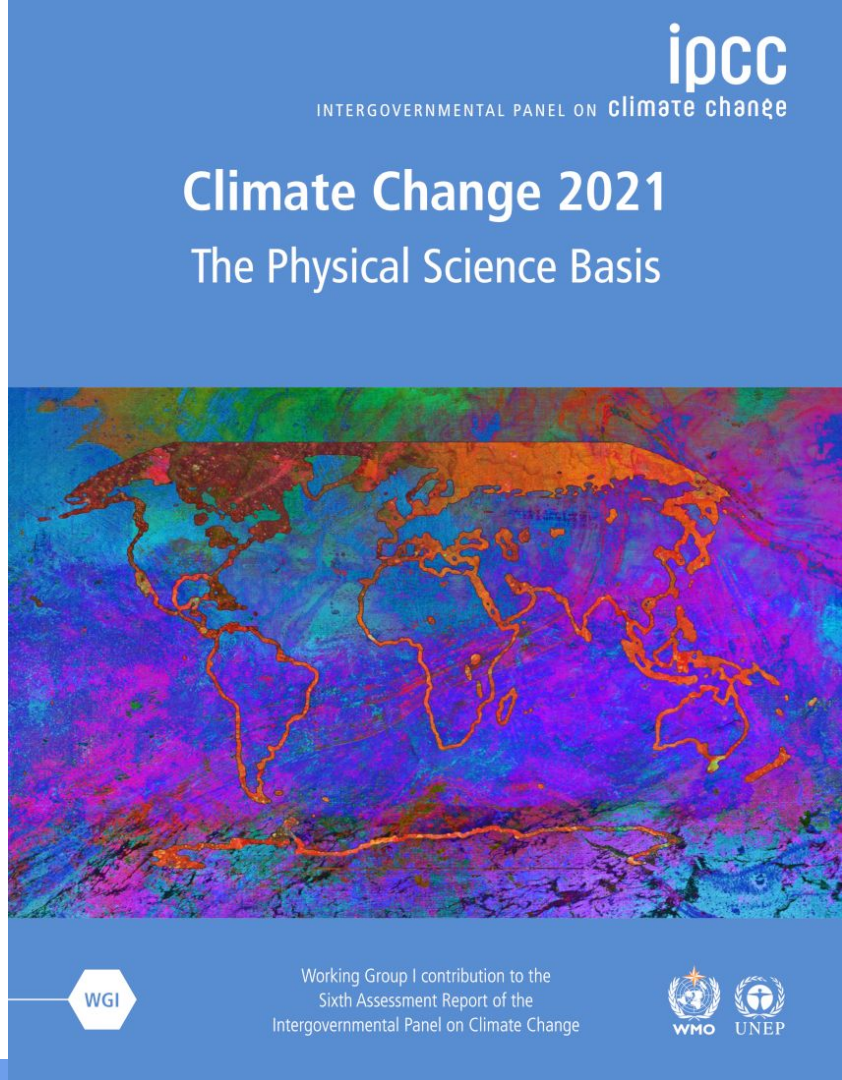
ESGF2 What is the Earth System Grid Federation?

- The **Earth System Grid Federation (ESGF)** is a globally distributed peer-to-peer network of data servers using a common set of protocols and interfaces to archive and distribute Earth system model (ESM) output
- ESM output data are used by scientists all over the world to investigate consequences of possible climate change scenarios and the resulting Earth system feedbacks



ESGF2 IPCC AR6 Released

- 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
- CMIP6 totals >20 PB
- We expect CMIP7 output, including high resolutions simulations and more ensembles, to total >100 PB
- We plan to expand Federation holdings by adding other Earth science data projects

CMIP6	11,307,181 total datasets 20,823.68 TB	5,400,359 distinct datasets 11,236.58 TB	5,906,822 replica datasets 9,587.1 TB
CORDEX	183,980 total datasets 1,391.12 TB	183,708 distinct datasets 1,390.56 TB	272 replica datasets 0.56 TB
CMIP5	201,129 total datasets 5,295.44 TB	52,163 distinct datasets 1,527.12 TB	148,966 replica datasets 3,768.32 TB
INPUT4MIPS	11,492 total datasets 19.91 TB	5,660 distinct datasets 9.97 TB	5,832 replica datasets 9.94 TB
OBS4MIPS	210 total datasets 0.2 TB	210 distinct datasets 0.2 TB	0 replica datasets 0 TB

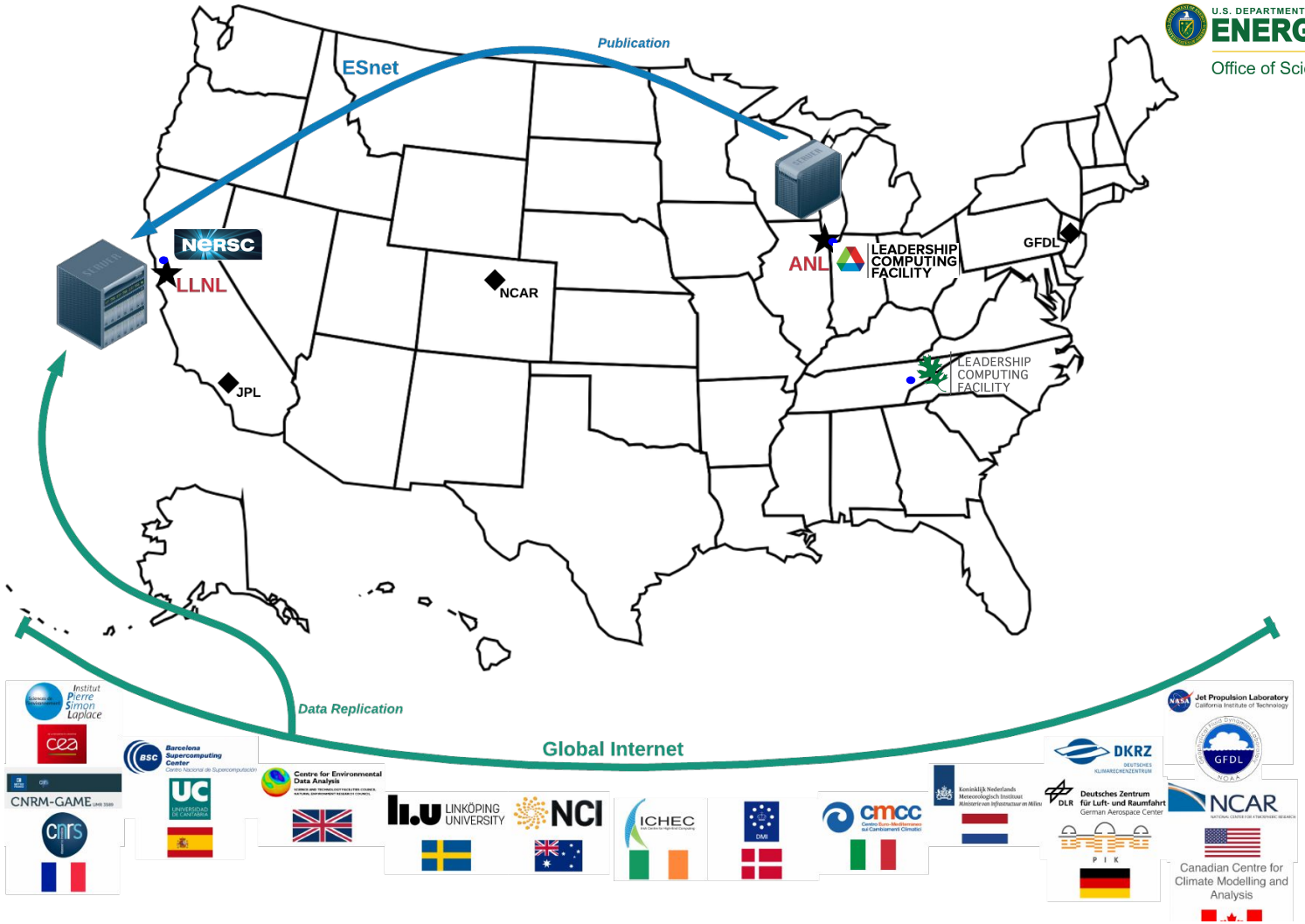
As of August 22, 2021

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 proposal was reviewed by panel of 8 scientists on August 30–31, 2021, and was **selected for funding** by the US Department of Energy in September
- In collaboration with the **ESGF Executive Committee**, we will develop and deploy a new architecture based on the *Future Architecture Roadmap*
- In addition, we 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**
- We will add a **Resource & Project Liaison** group and a **Science, User & Facility Advisory Board**; hold outreach activities; and offer a help desk/user support

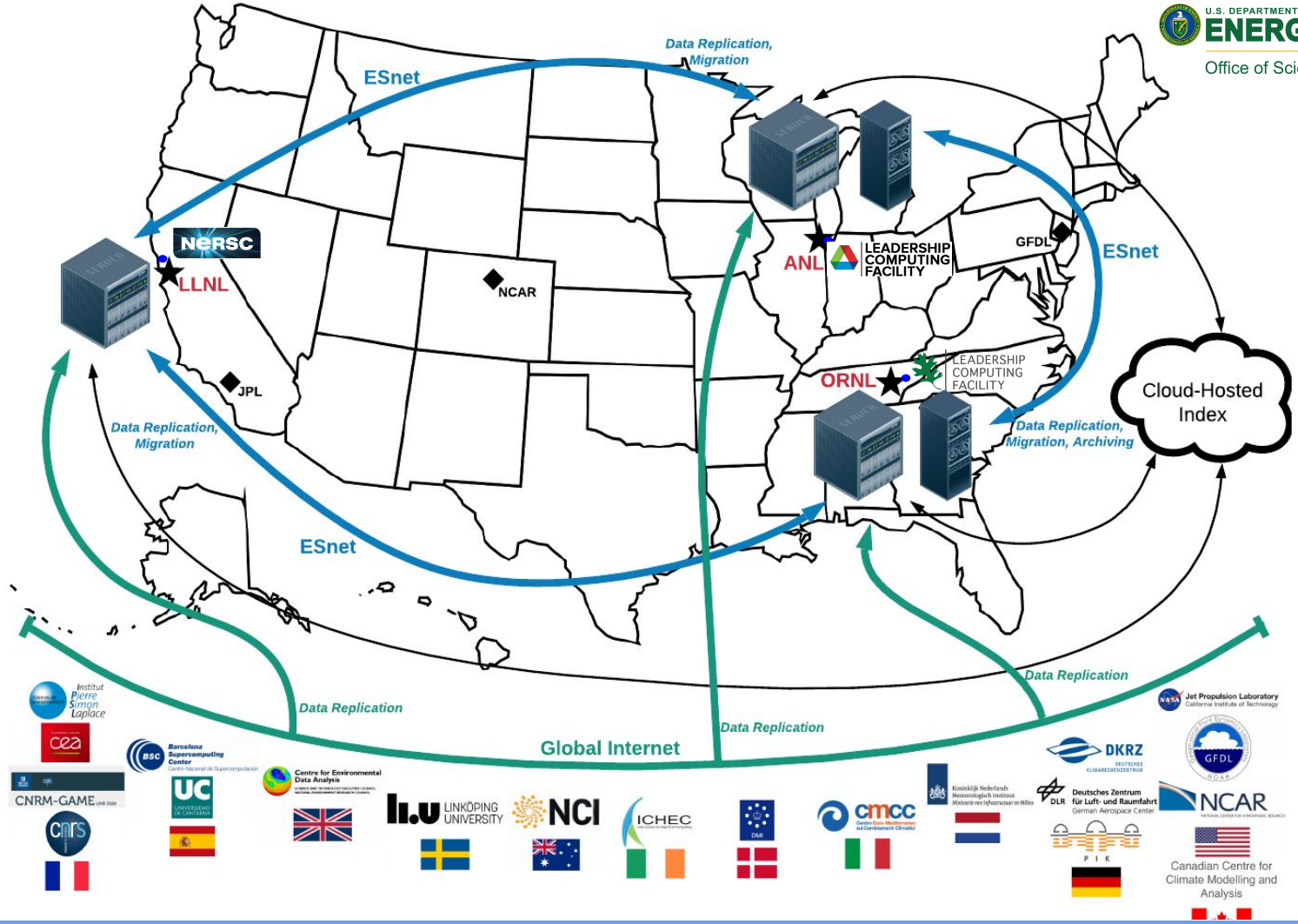
DOE's Current Earth System Grid Federation

- Primary server at LLNL
- Replicating data from the global Federation
- Independent data node at ANL



DOE's Next Generation Earth System Grid Federation

- Co-located at DOE's major computing facilities
- Replicating data from the global Federation
- Providing cloud indexing, automated migration, and tape archiving

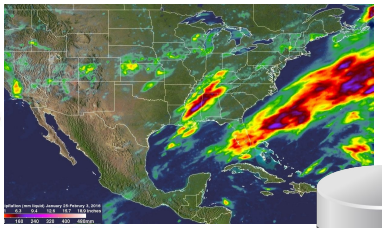


Design and implementation principles

- **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 Enabling a new level of research productivity

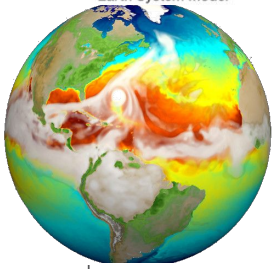
*Logging in with her **institutional credentials**, Samantha is presented with **new data, code, and papers** relevant to her current research. Intrigued by a new report on extreme precipitation events, she examines a **Jupyter notebook** that implements the method used. Wondering how this method would 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 federated storage, and they agree that a larger ensemble analysis is called for. ESGF2 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 for broader consumption**, along with the notebook used to produce and validate the results.*



ESGF2

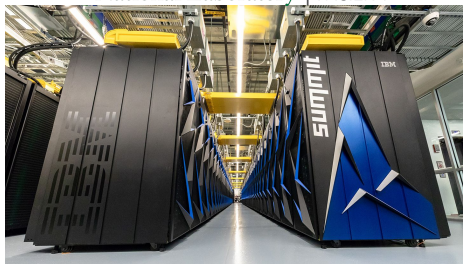
E3SM Energy Exascale Earth System Model

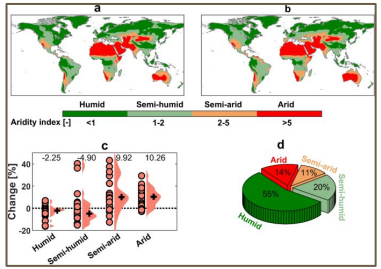
GitHub



jupyter

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY





Flood risk increases with water availability

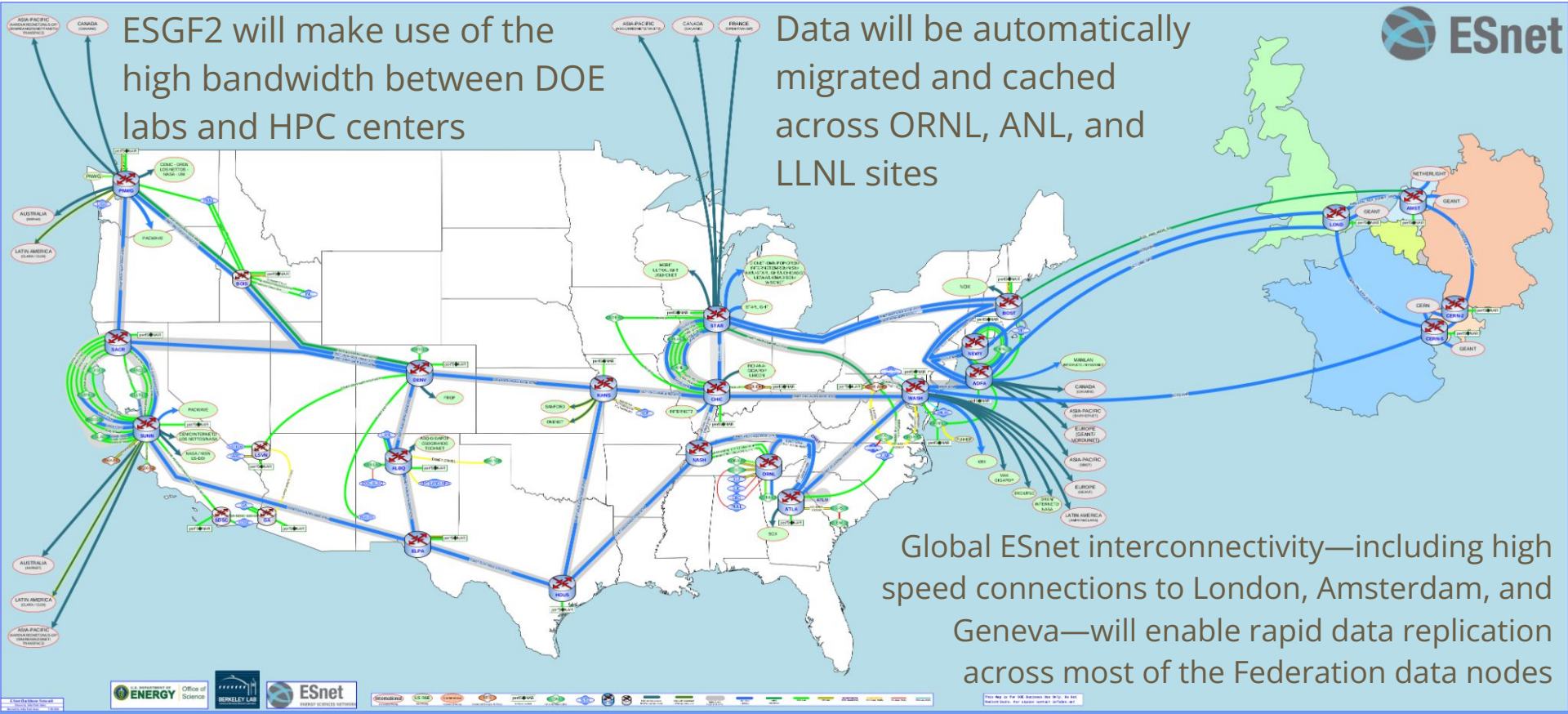
ESGF2 ESnet Global Connectivity

An ESnet representative is part of our Resource & Project Liaisons group

ESnet
ENERGY SCIENCES NETWORK

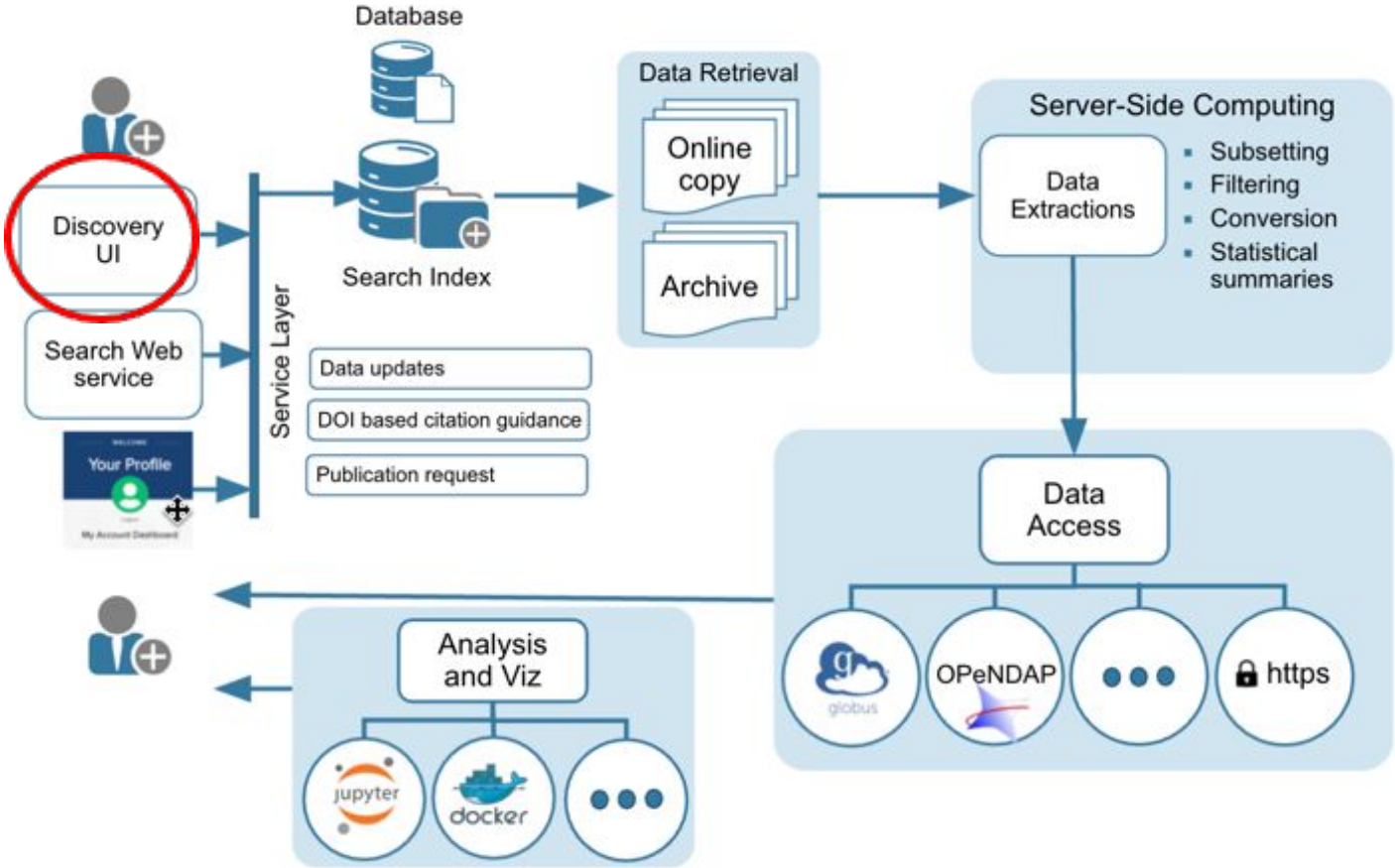
ESGF2 will make use of the high bandwidth between DOE labs and HPC centers

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

ESGF2 Data Discovery Platform: Architecture

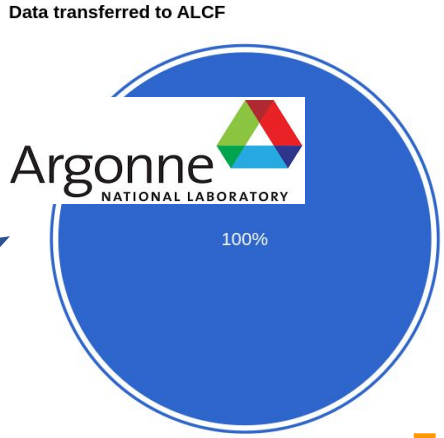


ESGF2 Outreach Activities

- Organize Webinars, Tutorials, and ESGF2 Bootcamps
  - Data management lessons learned
  - Ingest best practices
  - Data discovery and access
- Hackathons and Workshops
  - Data standards
  - Data node deployment
  - User compute resources
  - Hold at large relevant conferences, e.g., AGU Fall Meeting, EGU, and AMS Annual Meeting
- Organize and host an annual **ESGF Developer and User Conference**; next conference scheduled for January 18–20, 2023

ESGF2 ESGF-US Failsafe Data Replication

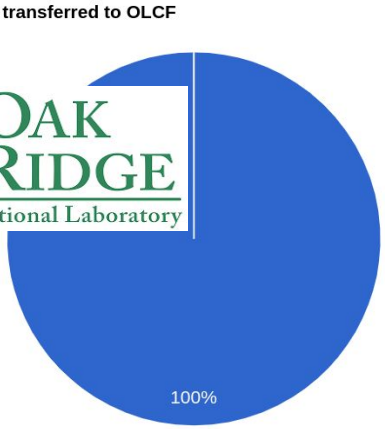
- **In the US, LLNL operates the primary ESGF node**, which replicates much of the CMIP6 and related model output from around the globe
- Since the data at LLNL are contained only on spinning disk, we decided to replicate the **entire ~7.5 PB collection of data** to Argonne National Laboratory (ANL) 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*)



1.5 GB/s

Lawrence Livermore National Laboratory

4 to 6 GB/s



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

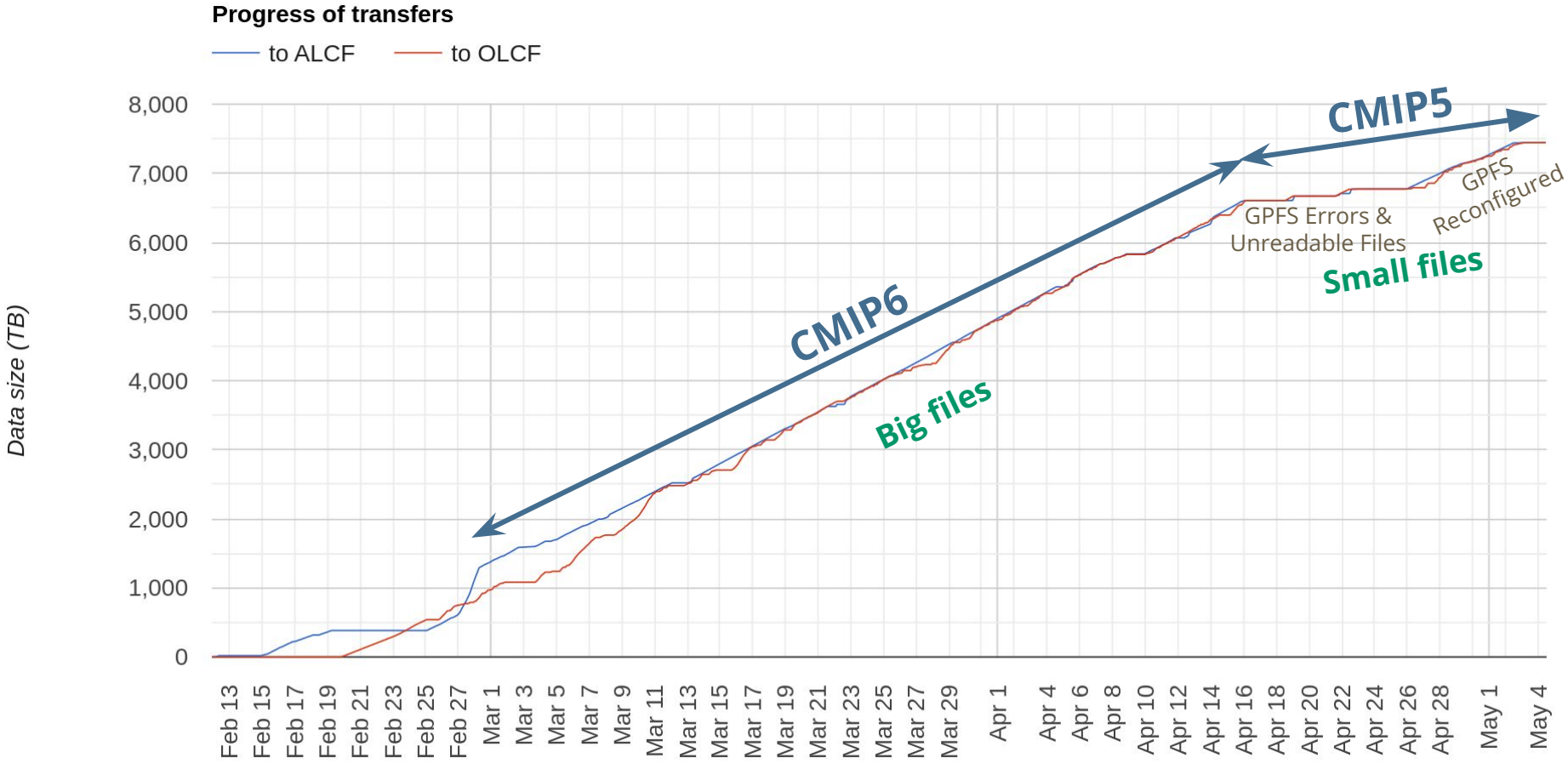
ESnet ENERGY SCIENCES NETWORK

globus

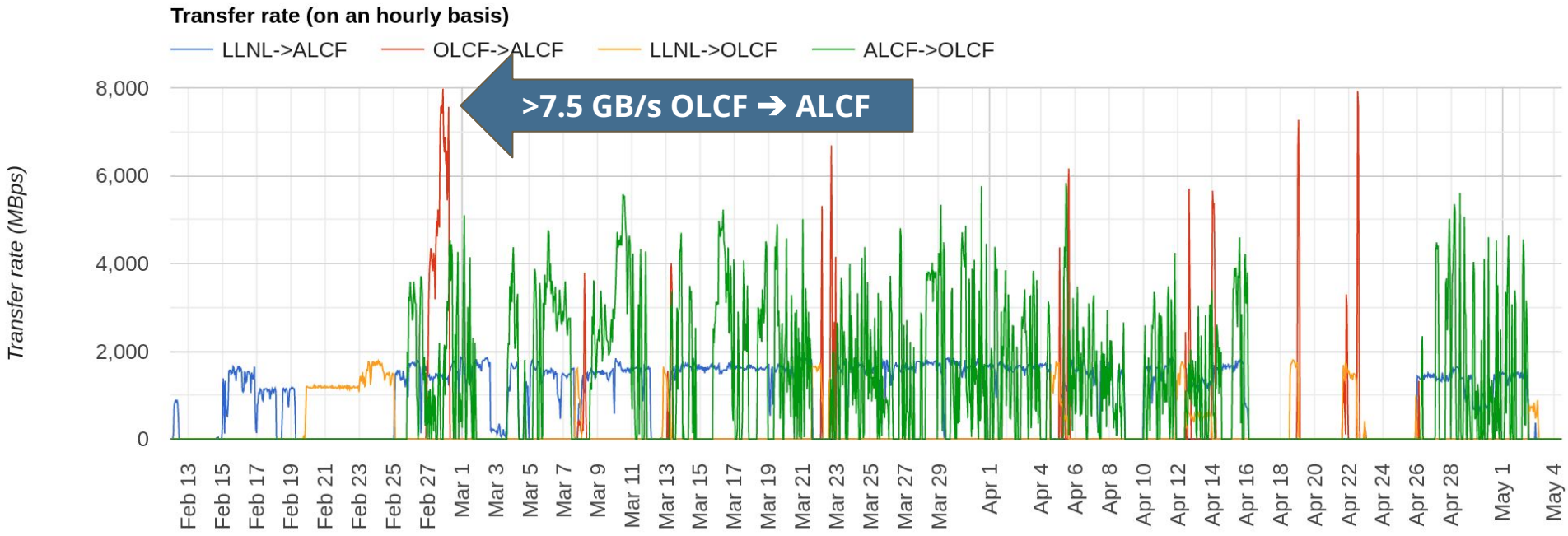
https://dashboard.globus.org/esgf

As of May 4, 2022

ESGF2 Cumulative Data Transferred Over Time



ESGF2 Transfer Rates Over Time



ESGF2 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
- The **Globus platform** is expected to provide many of the central services of the ESGF2-US data backplane in the new deployment
- We used Globus to make two redundant copies of the **7.5 PB of ESGF data** via ESnet in less than 3 months