International Consortium Developing the Next Generation Earth System Grid Federation (ESGF) Distributed Data Infrastructure

Forrest M. Hoffman (ORNL),
Ian Foster (ANL), and Sasha Ames (LLNL)

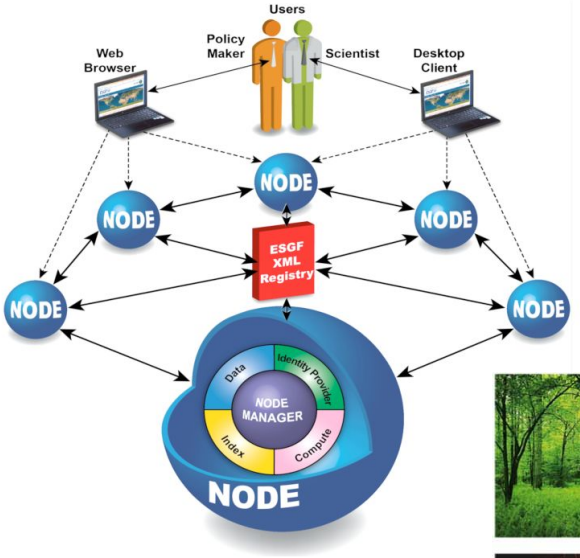
Interagency Arctic Research Policy Committee (IARPC) Collaborations
Meeting of MOMP, Data Management, and Modelers Teams
March 20, 2025

ESGF2-US Project is supported by the DOE Biological & Environmental Research (BER) Data Management Program in the Earth & Environmental Systems Sciences Division (EESSD) led by Dr. Jay Hnilo

What is the Earth System Grid Federation?

- **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 & interfaces to archive and distribute Earth system model output and related input, observational, and reanalysis data
- **Open Science data** are used by scientists all over the world to investigate Earth system variability and feedbacks and to inform assessment reports

ESGF Conceptual Diagram



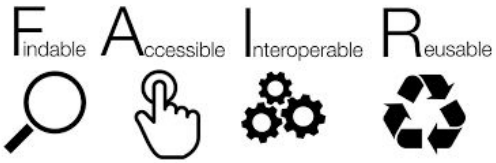
ESGF Earth System Grid Federation

Model data from ESGF are used to understand key Earth system processes and feedbacks

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

ESGF Holdings are Open and Large

- CMIP5 totals >1.5 PB (>5 PB including replicas)
- CMIP6 totals >15.9 PB (>27 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* (mostly)

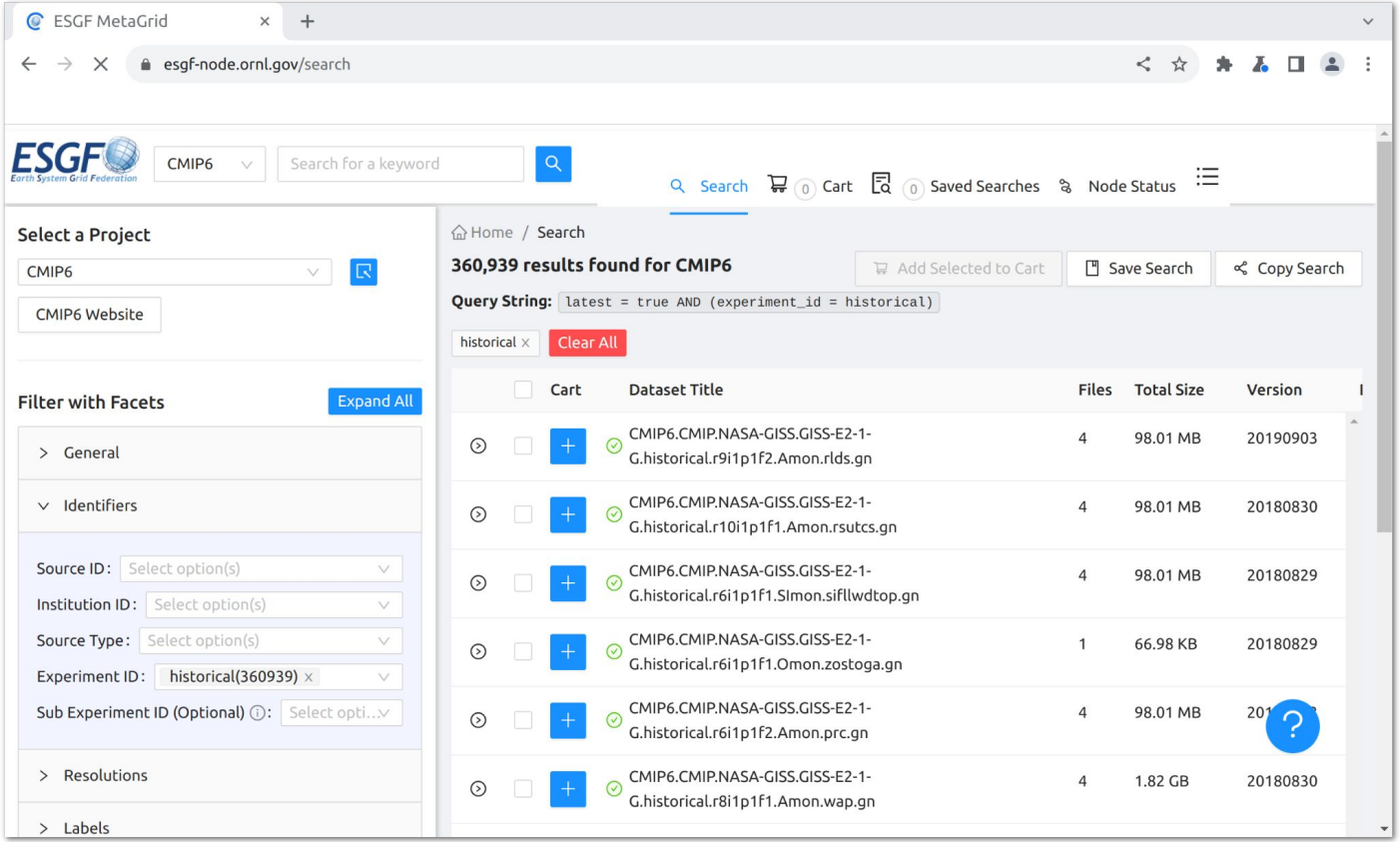


CMIP6: 14,780,267 total datasets, 27,422.79 TB	CMIP6: 7,585,457 distinct datasets, 15,929.59 TB	CMIP6: 7,194,810 replica datasets, 11,493.2 TB
CORDEX: 187,785 total datasets, 1,473.33 TB	CORDEX: 187,513 distinct datasets, 1,472.77 TB	CORDEX: 272 replica datasets, 0.56 TB
CMIP5: 201,130 total datasets, 5,293.61 TB	CMIP5: 52,163 distinct datasets, 1,525.07 TB	CMIP5: 148,967 replica datasets, 3,768.55 TB
INPUT4MIPS: 5,871 total datasets, 10.84 TB	INPUT4MIPS: 0 distinct datasets, 0 TB	INPUT4MIPS: 5,871 replica datasets, 10.84 TB
OBS4MIPS: 126 total datasets, 0.2 TB	OBS4MIPS: 108 distinct datasets, 0.2 TB	OBS4MIPS: 18 replica datasets, 0.01 TB

As of June 20, 2024

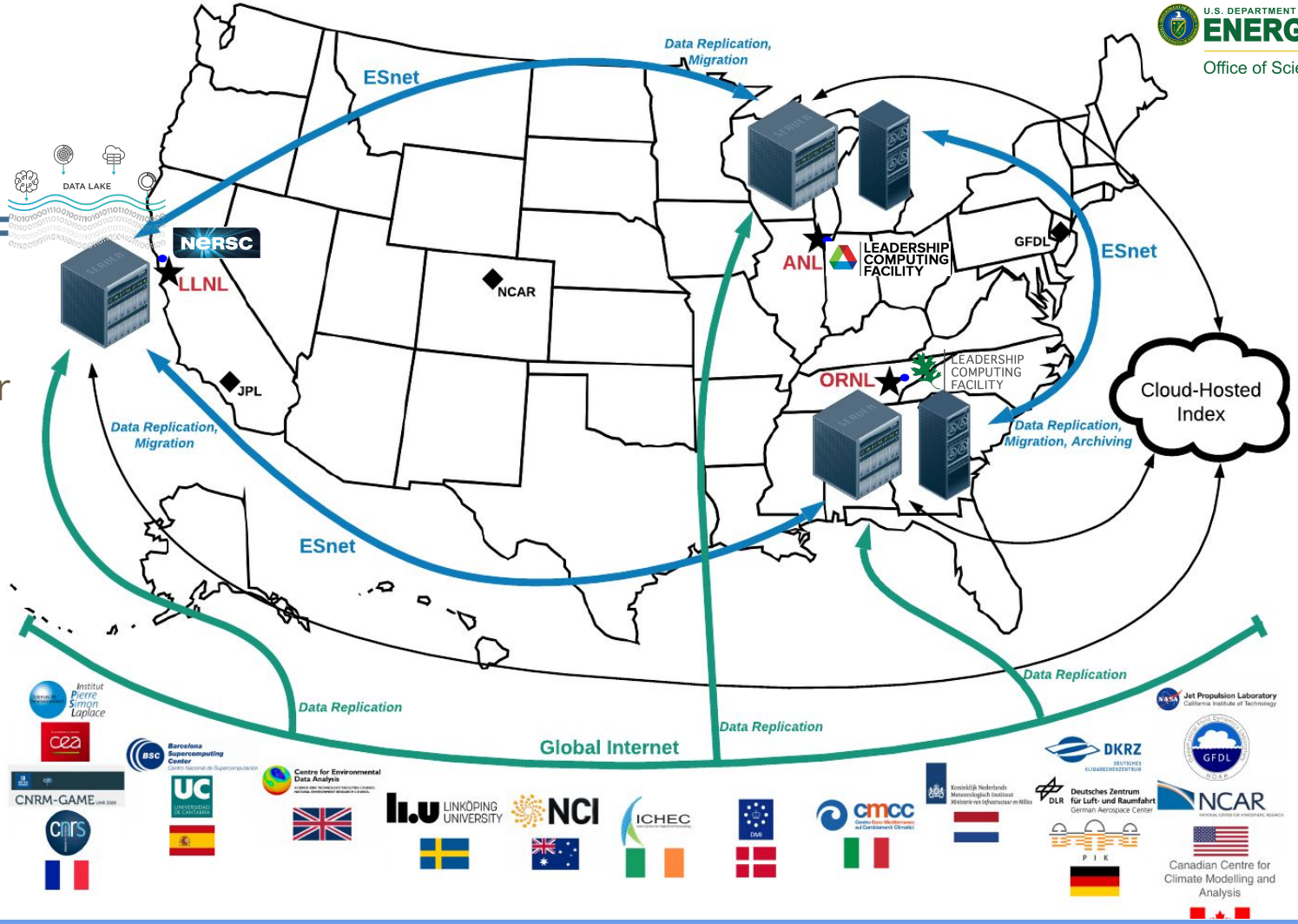
ESGF2 US Metagrid Data Search Interface

- A redesigned faceted search user interface, called **Metagrid**, replaces the old interface and adds new features
- Offers **shopping cart** and ability to **save & share searches**
- Will soon provide **Globus integration** for fast unattended data transfers



DOE's Next Generation ESGF

- As many as 3 nodes located at DOE's major computing facilities
- Replicating data from the worldwide Federation
- Providing scalable cloud indexing and tape archiving



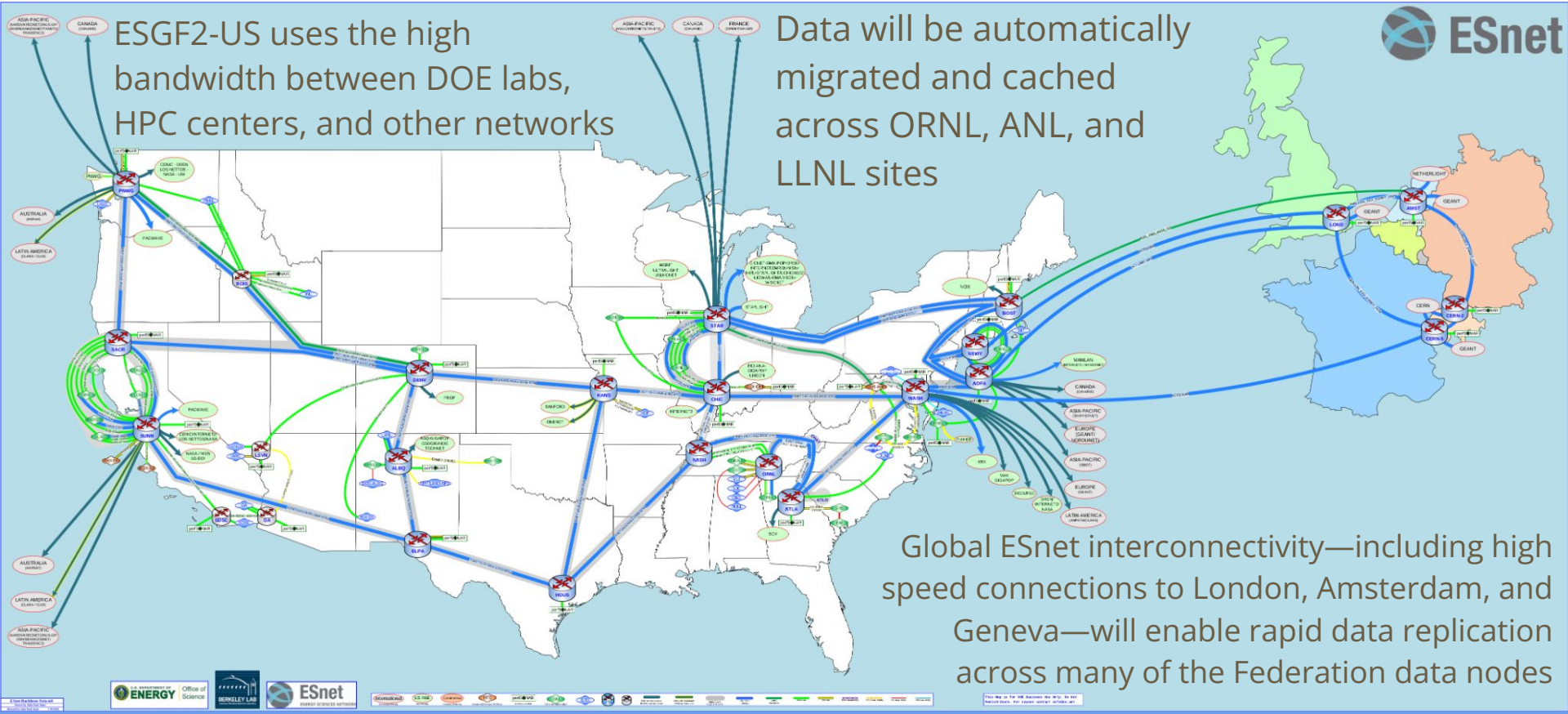
ESGF2 US ESnet Global Connectivity

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

ESnet
ENERGY SCIENCES NETWORK

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

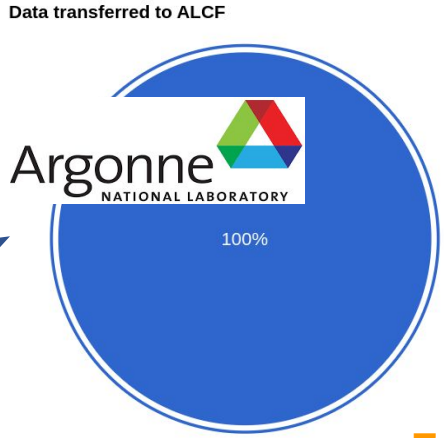
Data will be automatically migrated and cached across ORNL, ANL, and LLNL sites



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 Failsafe Data Replication

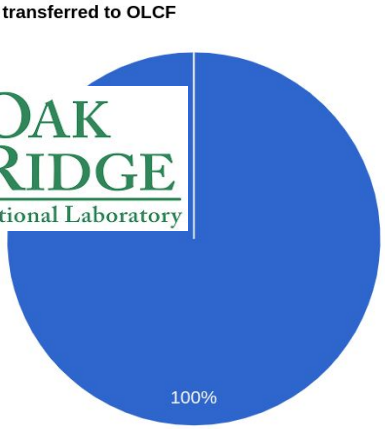
- **In the US, LLNL operates the primary ESGF node**, which replicates much of the CMIP6 and related model output from around the globe
- Since the data at LLNL are contained only on spinning disk, we decided to replicate the **entire ~7.5 PB collection of data** to Argonne National Laboratory (ANL)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

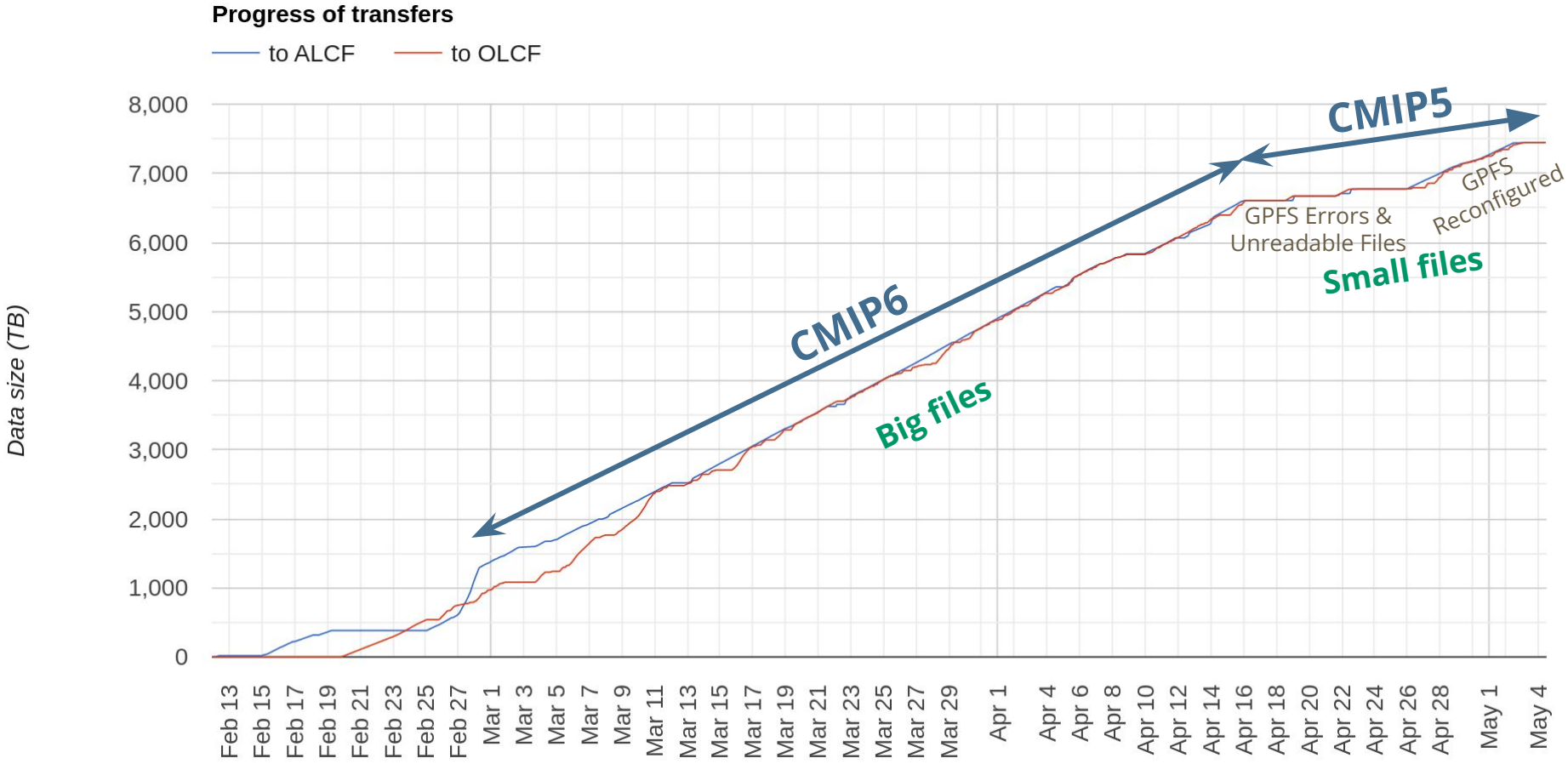
ESnet ENERGY SCIENCES NETWORK

globus

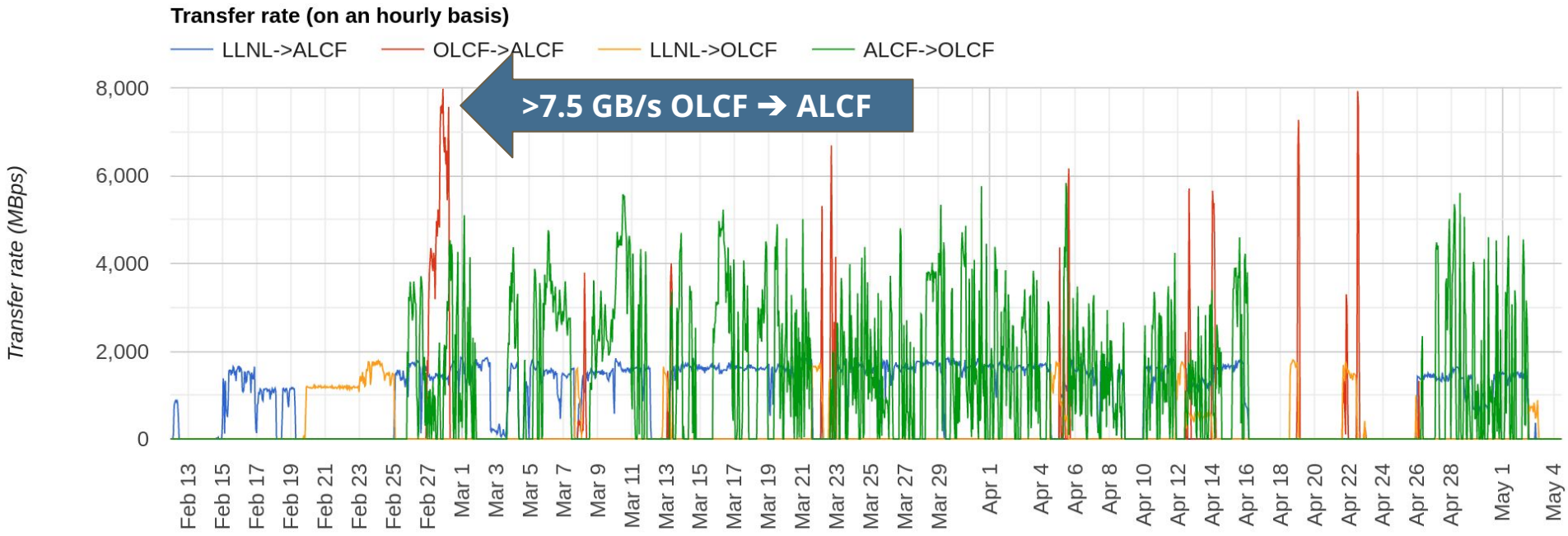
https://dashboard.globus.org/esgf

As of May 4, 2022

ESGF2 US Cumulative Data Transferred Over Time

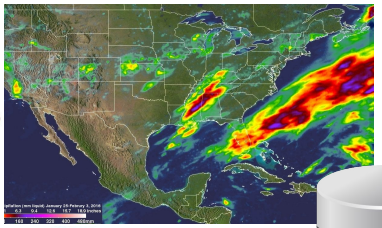


ESGF2 US Transfer Rates Over Time



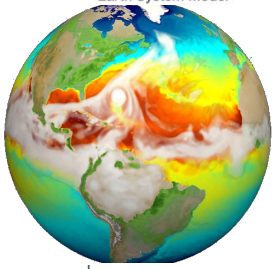
ESGF2 US Enabling a new level of research productivity

*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.*



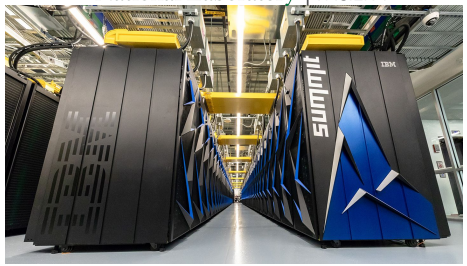
E3SM
Energy Exascale Earth System Model

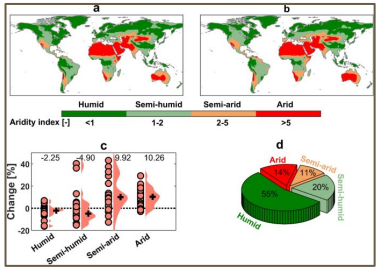
GitHub



jupyter

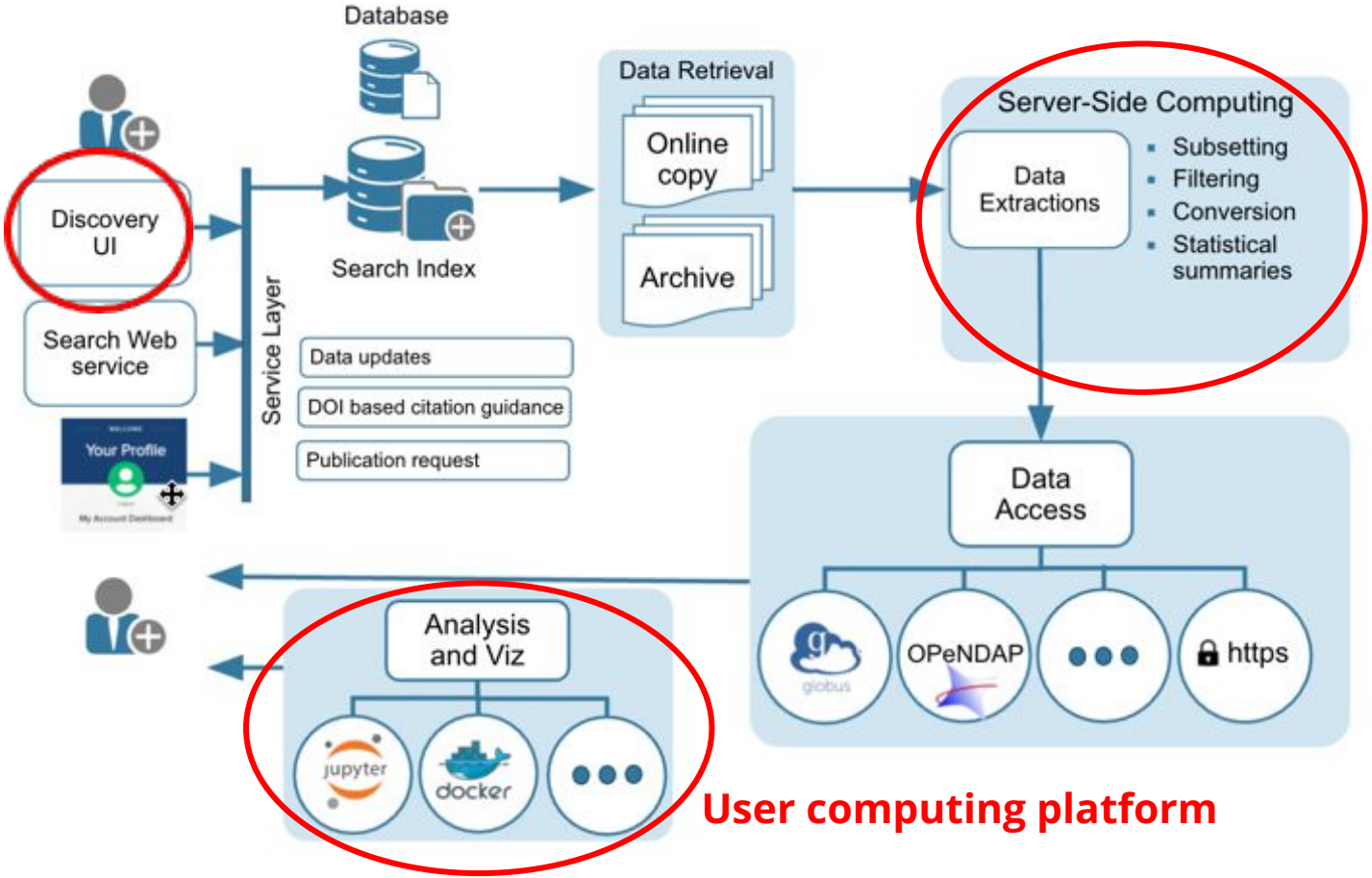
OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY





Flood risk increases with water availability

ESGF 2 US Data Discovery Platform: Architecture



Friendly user interface

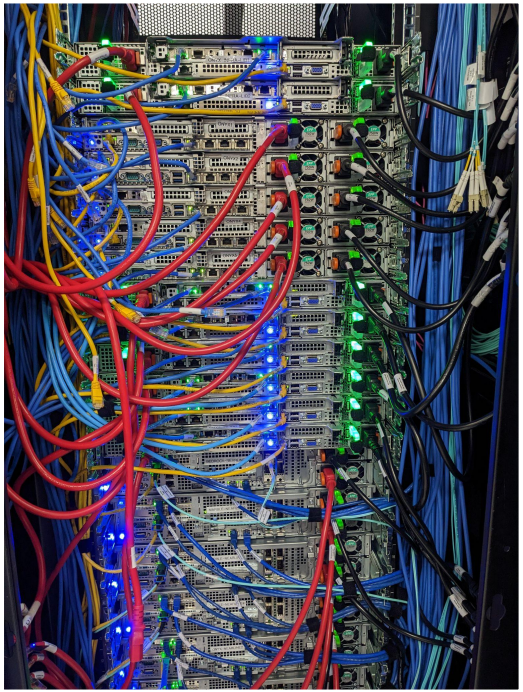
Server-side computing platform

User computing platform

ESGF2 US Data and Index Nodes Deployed at ORNL

- Containerized server software deployed on the shared Onyx cluster is serving 8 PB of Coupled Model Intercomparison Project (CMIP5 and CMIP6) data at ORNL
- Data are stored on the new Themis hierarchical storage platform, providing on-disk copy for fast access to frequently used data and backup copies on two tapes for all data
- Hardware investment at ORNL has been in storage capacity (fully operational)
  - 15 PB of disk
  - 30 PB of tape (for redundant backup)

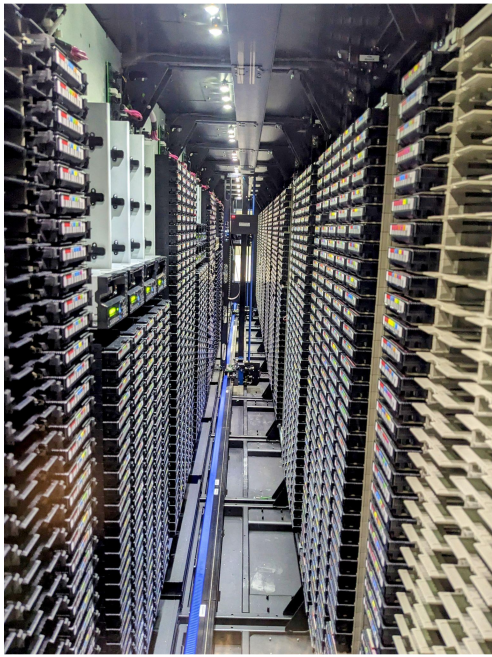
Delivered ahead of schedule and under budget!



The Onyx cluster hosts the ESGF containerized data & index nodes

Data and services reside in the Open Network Enclave of NCCS to provide fast and open access to data

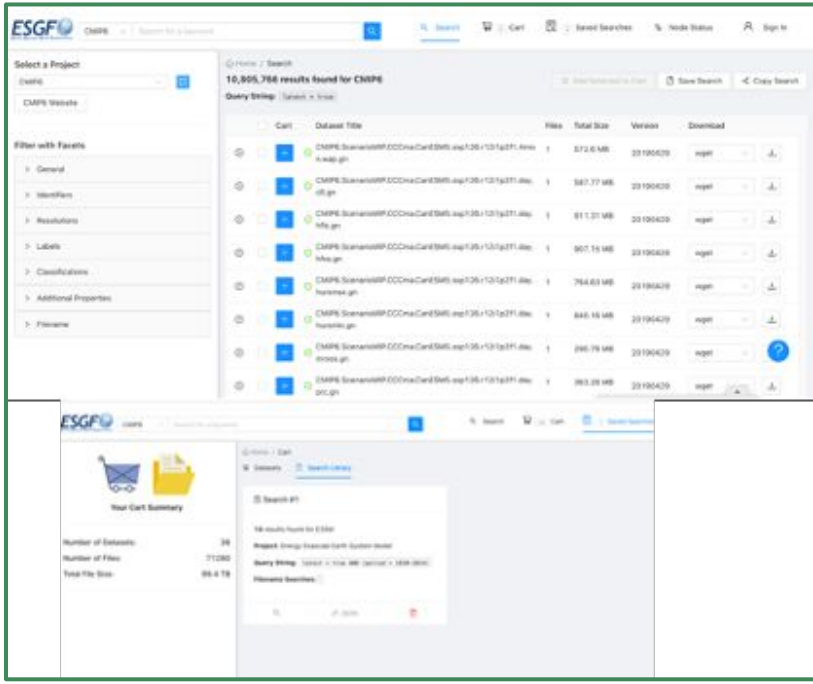
In partnership with the ORNL National Center for Computational Sciences (NCCS)



Expandable tape subsystem of the Themis storage system

ESGF2 US Metagrid Enhances ESGF Search

- New **Metagrid faceted search user interface**, developed at LLNL on popular React Javascript framework, deployed at ORNL, LLNL and ANL
- Offers new features, including a **shopping cart**, ability to **save and share searches**, integration with **Globus authentication & transfer** and a search page **tour & support dialog**
- User experience enhancements make it **faster and easier** to discover published data
- **Globus integration** offers faster and more reliable data access
- Will be deployed internationally across the Federation by mid-2024



The Metagrid Web Interface for ESGF search is a completely redesigned interface from CoG. It features a familiar faceted search and a new capability to save searches.

ESGF2 US Outreach Activities

- Organize **Webinars, Tutorials, and Bootcamps**
  - Data management lessons learned, ingest best practices
  - Data discovery and access, analysis frameworks and tools

➜ **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, EGU, AMS

➜ **Open ESGF Workshop at AGU 2022 (Chicago)**

➜ **Open ESGF Workshop & Tutorial at AGU 2023 (San Francisco)**

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

➜ **Ninth ESGF Developer and User Dual-Hybrid Conference was held January 18–20, 2023 at ORNL and Toulouse**

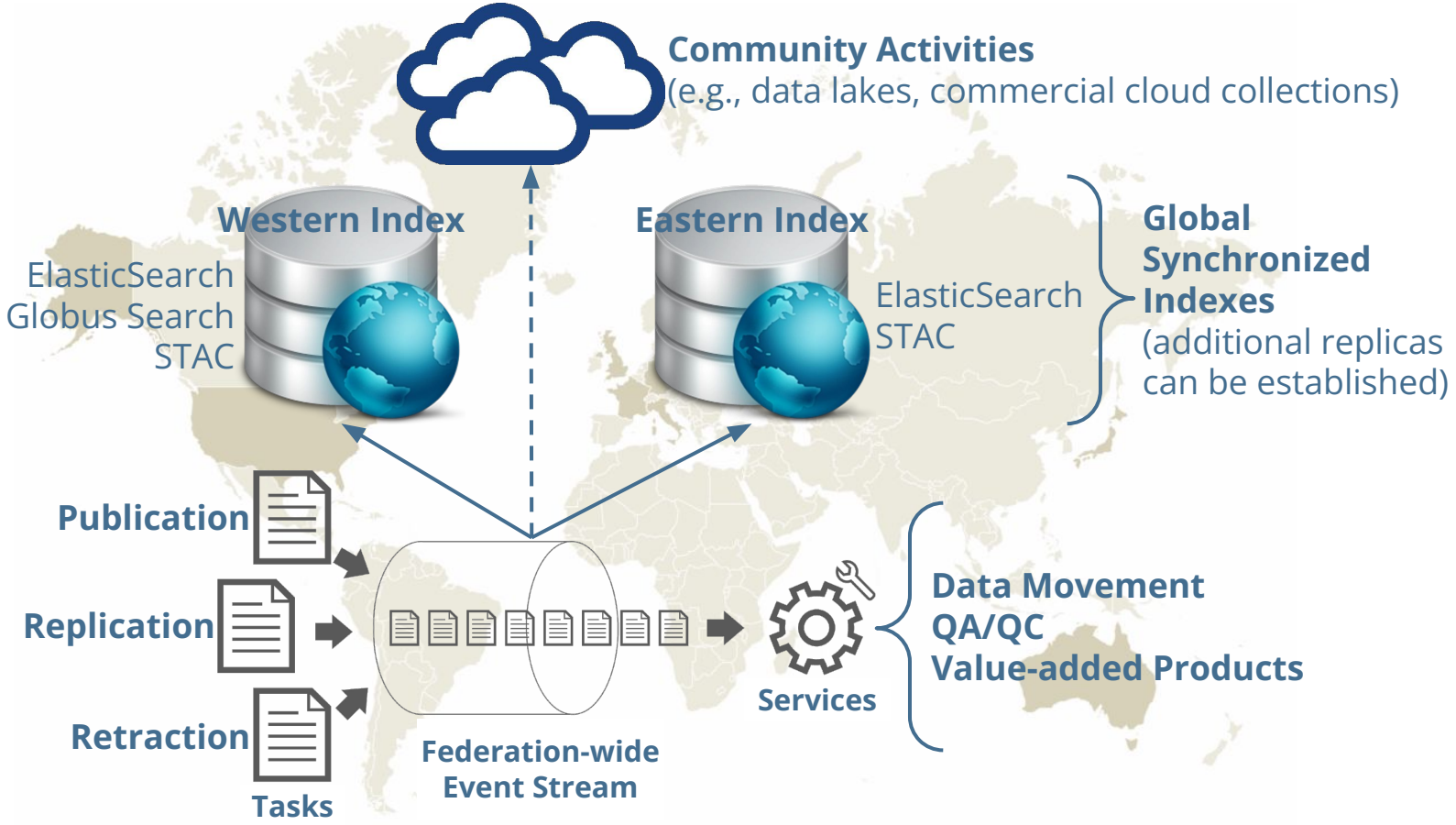
➜ **Tenth ESGF Developer and User Conference scheduled for Rockville, MD, on April 23–26, 2024**

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

Tenth ESGF Conference

- Held 23–26 April 2024 in Rockville, Maryland
- John Dunne joined the meeting to share CMIP priorities and current CMIP timeline
- ~50 in-person attendees from 8 countries (Australia, France, Germany, Italy, Japan, Sweden, United Kingdom, USA)
- ~69 virtual registrants from 18 countries
- **Primary objectives** of conference were to
  - **Share all current development activities** across the Federation
  - **Develop a roadmap** for collaborative activities necessary to deploy operational ESGF infrastructure to support CMIP AR7 Fast Track

ESGF2 US Major Accomplishment: New Index Strategy



ESGF 2 US Summary of Integration Activities

- All **ESGF development is being performed collaboratively** with Federation partners
- **New data projects** for downscaled projections (LOCA2, STAR-ESDM) were added; we will add large-scale AI/ML data, large ensembles simulations and intercomparisons
- **User computing** approaches initiated in the commercial cloud and deployed through on-premise cloud infrastructure will enable computing near the data
- Specific **integration activities**:
  - **Sharing data indexes** across DOE-BER platforms (ARM Data Center, ESS-DIVE, etc.)
  - Unifying on **Federated authentication** (*Globus Auth*) to simplify data access and enable cross-platform/cross-facility data access and analysis
  - **Integrating software stacks** for data access, analysis, and visualization for Jupyter
  - A few global **scalable data index** and search instances (*Globus Search*)
  - **Managed automation** of data publishing workflows (*Globus Flows*)
  - **Server-side computing** spawned by web or Jupyter/Python (*Web Processing Service* and *Globus Compute*) for generating value-added products and subsetting & summarizing data across platforms and facilities