The Earth System Grid Federation:
Management of Distributed Data

D3 Workshop
September 25, 2019

Sasha Ames, CASC/LLNL
ames4@llnl.gov

LLNL-PRES-790557
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

ESGF Overview

ESGF Earth System Grid Federation

Motivation

- Scientific progress in the **climate realm** is critically dependent on the availability of a **reliable** infrastructure for data access and management
- Next-generation data ecosystems must ensure that the scientific investigation is completely **transparent, collaborative**, and **reproducible**, which are crucial attributes, given the field's high visibility and direct impact on **climate** research

Approach

- The ESGF's approach is to develop a close **collaboration** between interagency partners from **disparate** domains to enable a data ecosystem that can support a broad variety of data and disciplines

Major federated ESGF worldwide sites.

Impact

- 18 highly visible national and international climate data products, including E3SM, CMIP3,5 and 6 obs4MIPs (**observations**) input4MIPs ("**forcings**" for CMIP simulations)
- >5 PB of distributed community data, expected to expand beyond >40 PB
- 33 registered nodes at climate research centers spread across 21 countries
- >25,000 registered users
- 1,000's of peer-reviewed publications prepared using the ESGF data and resource

Reference: Dean Williams, V. Balaji, Luca Cinquini, Sébastien Denvil, Daniel Duffy, Ben Evans, Robert Ferraro, Rose Hansen, Michael Lautenschlager, Claire Trenham, "A Global Repository for Planet-Sized Experiments and Observations," Bulletin of the American Meteorological Society, June 2016, DOI: http://dx.doi.org/10.1175/BAMS-D-15-00132.1.

ESGF is….

- An open-source software-stack
  - “Homegrown” and third-party components
- Operational collaboration of services (data and computational) based on federated architecture
  - “Federation” of sites
- Software development collaboration
  - “Federation” of development groups
- DOE R&D Project

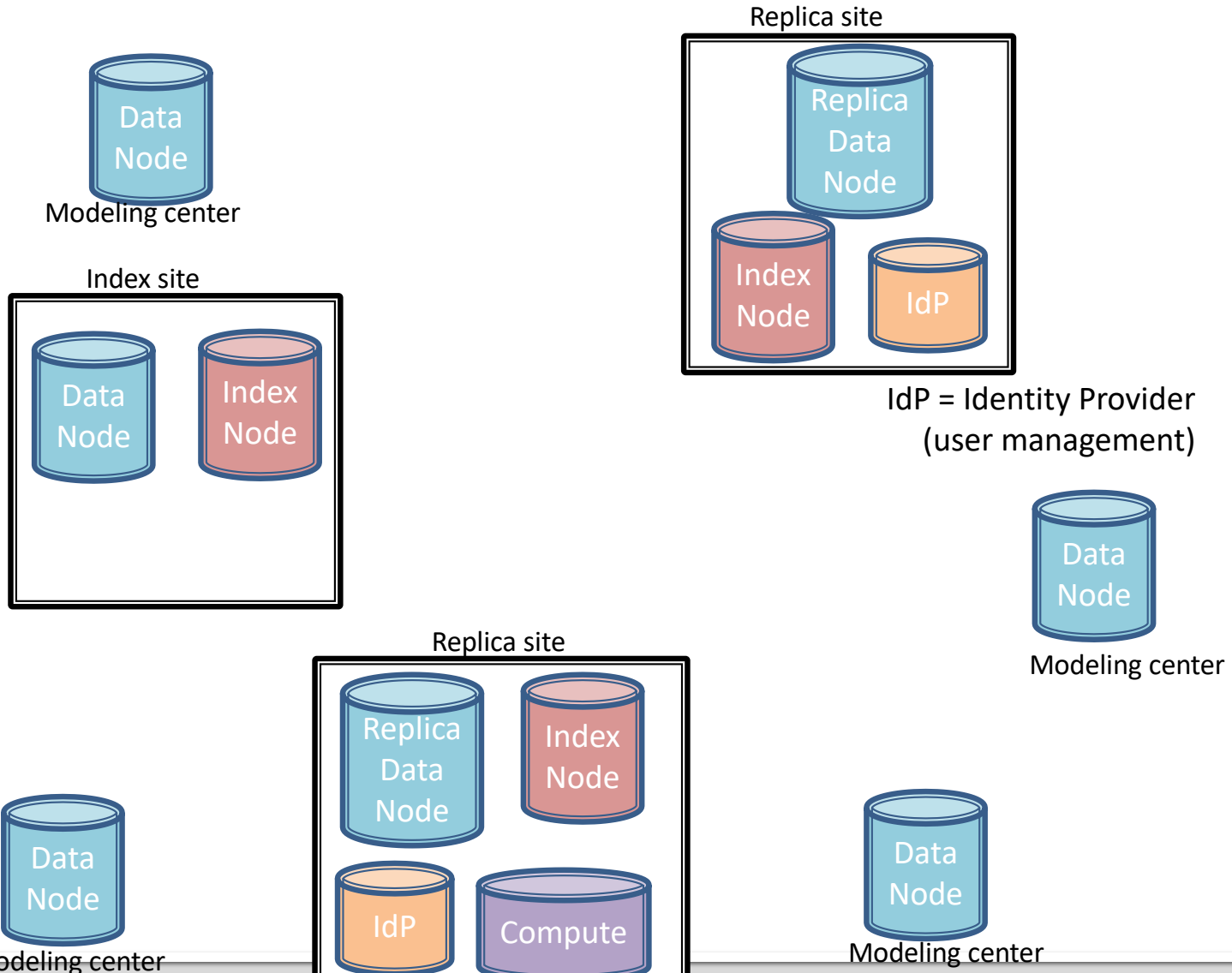
ESGF has led data archiving for the Coupled Model Intercomparison Project (CMIP) since its conception

- Gathering and sharing of climate data is a key effort of CMIP, the worldwide standard experimental protocol for studying general circulation model output
- This climate modeling research requires enormous scientific and computational resources that involves over 88 models and spans more than 20 countries
- The World Climate Research Program (WCRP) serves as the primary coordinating body for this research activity
- The WCRP Working Group on Coupled Modeling (WGCM) relies on the ESGF to support these activities by coordinating and maintaining the distributed petabyte data archive
- CMIP simulation model runs are key components of periodic assessments by the Intergovernmental Panel on Climate Change (IPCC).

Use cases for data sharing / access

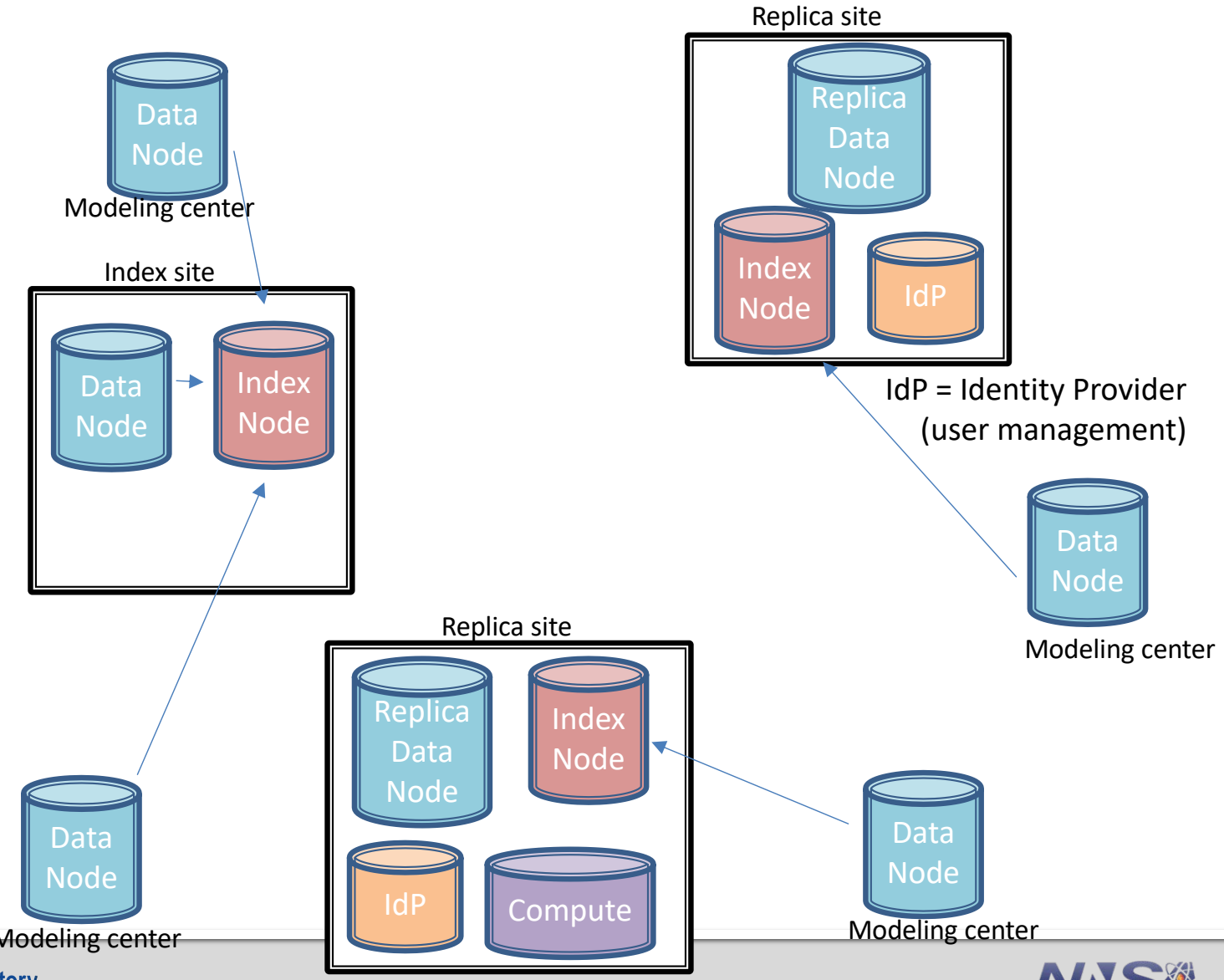
- Continued model development
- Basic science questions: what can we learn about how these complex systems work?
- Applied science questions: impacts of future climate projections

Federation Design



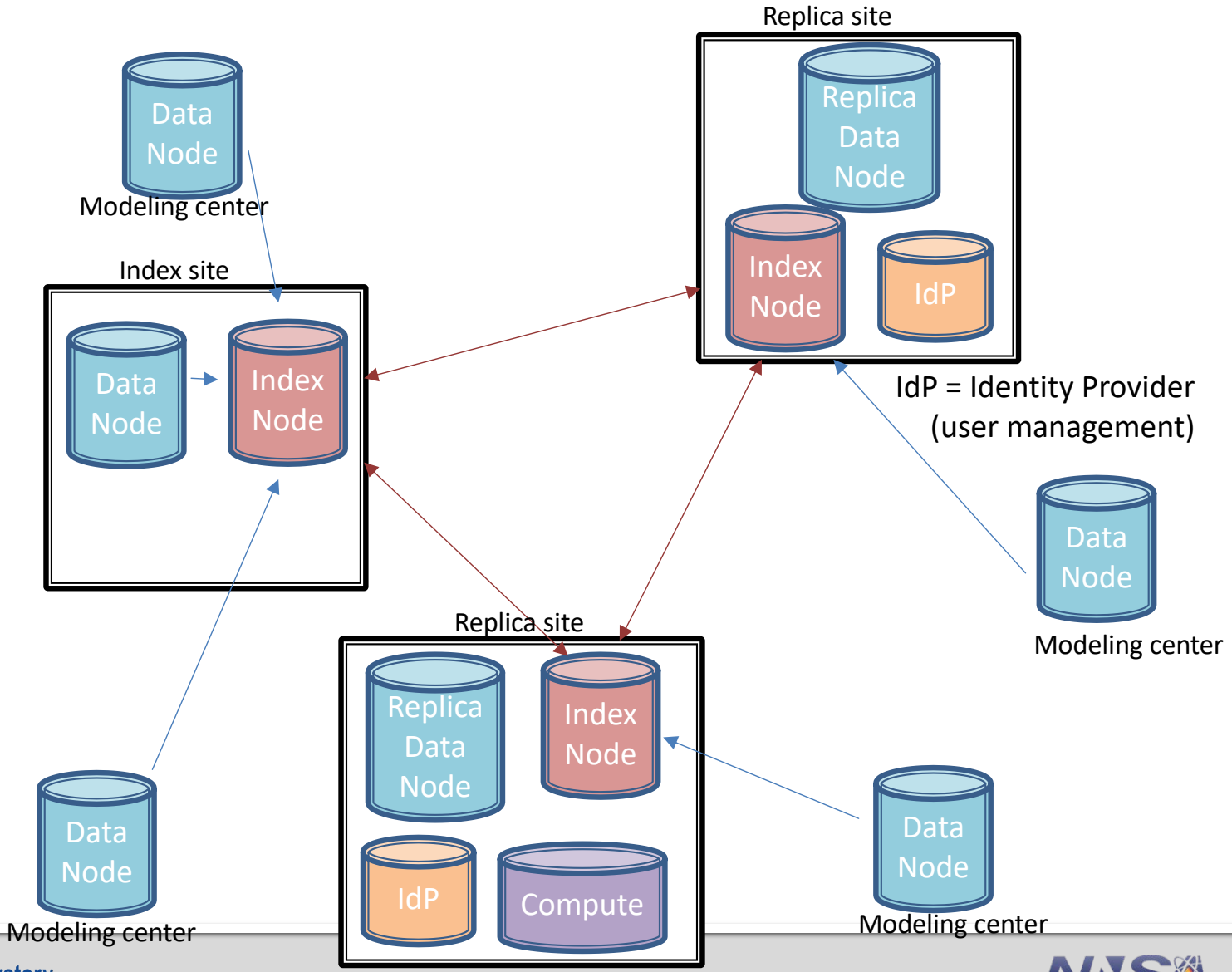
Federation Design

1. **Publication**: metadata and data locations stored on index nodes



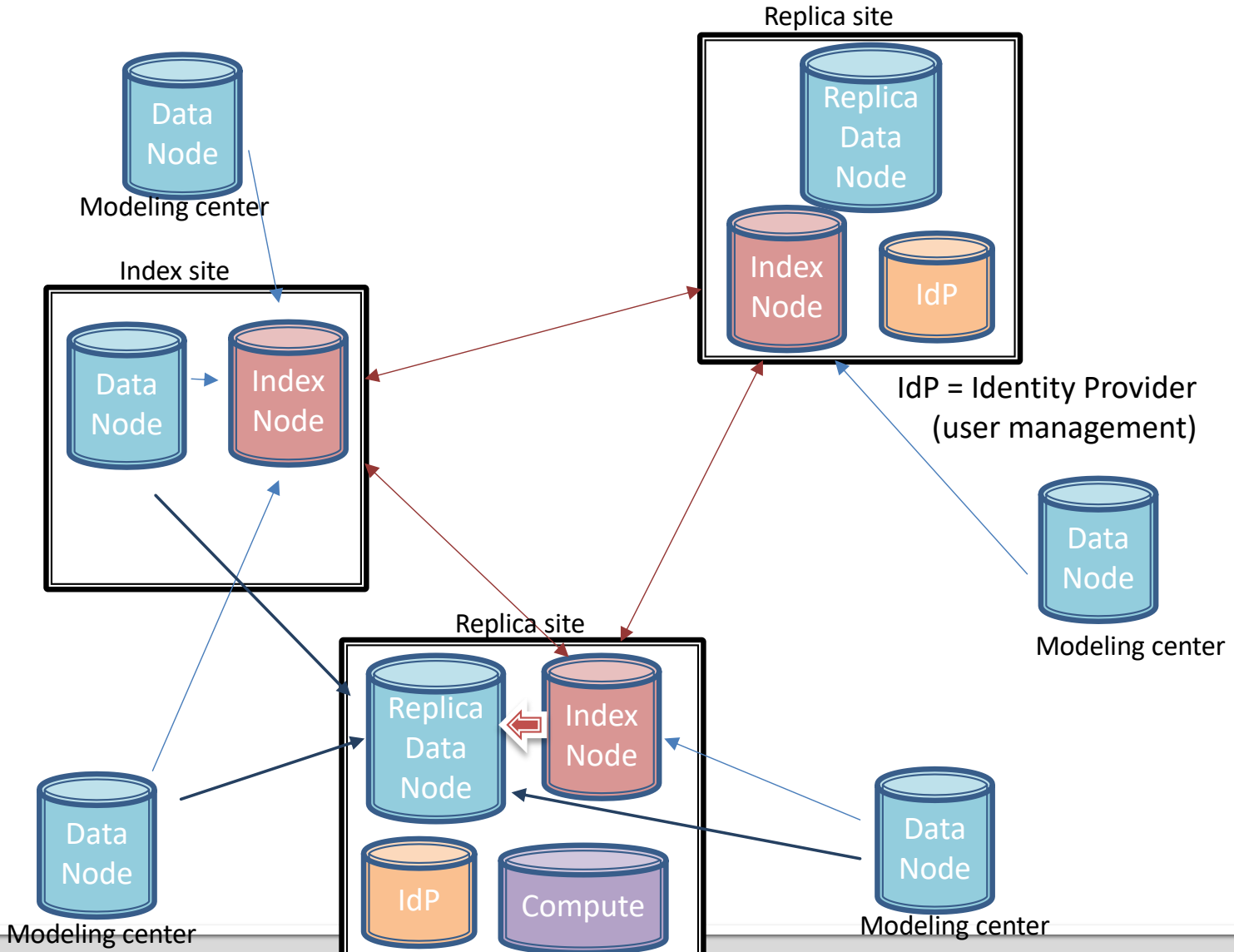
Federation Design

1. **Publication**: metadata and data locations stored on index nodes
2. **Index replication** across index sites



Federation Design

1. **Publication**: metadata and data locations stored on index nodes
2. **Index replication** across index sites
3. **Data Replication** informed by metadata in indexes



ESGF Node Software Components Architecture

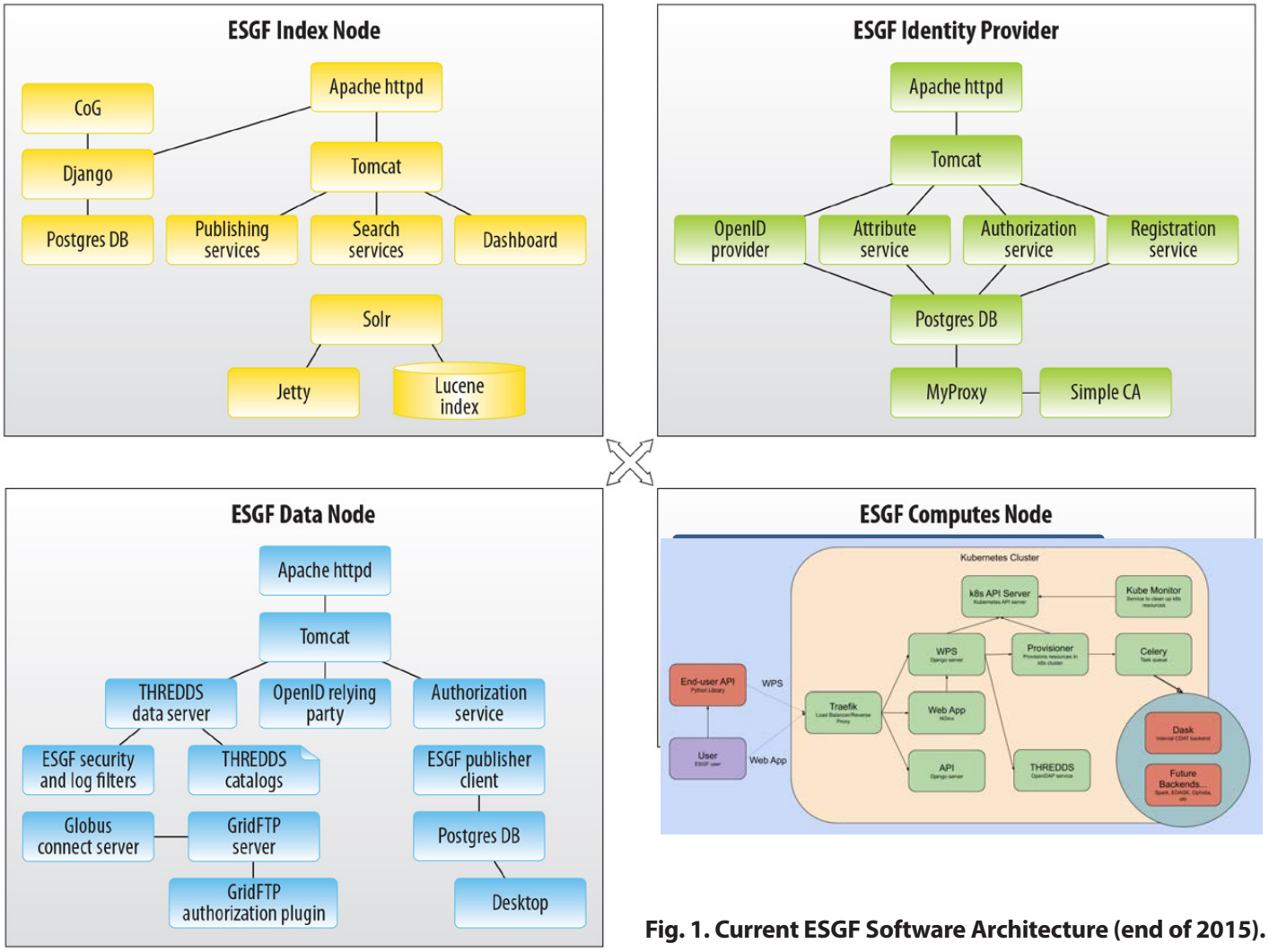


Fig. 1. Current ESGF Software Architecture (end of 2015).

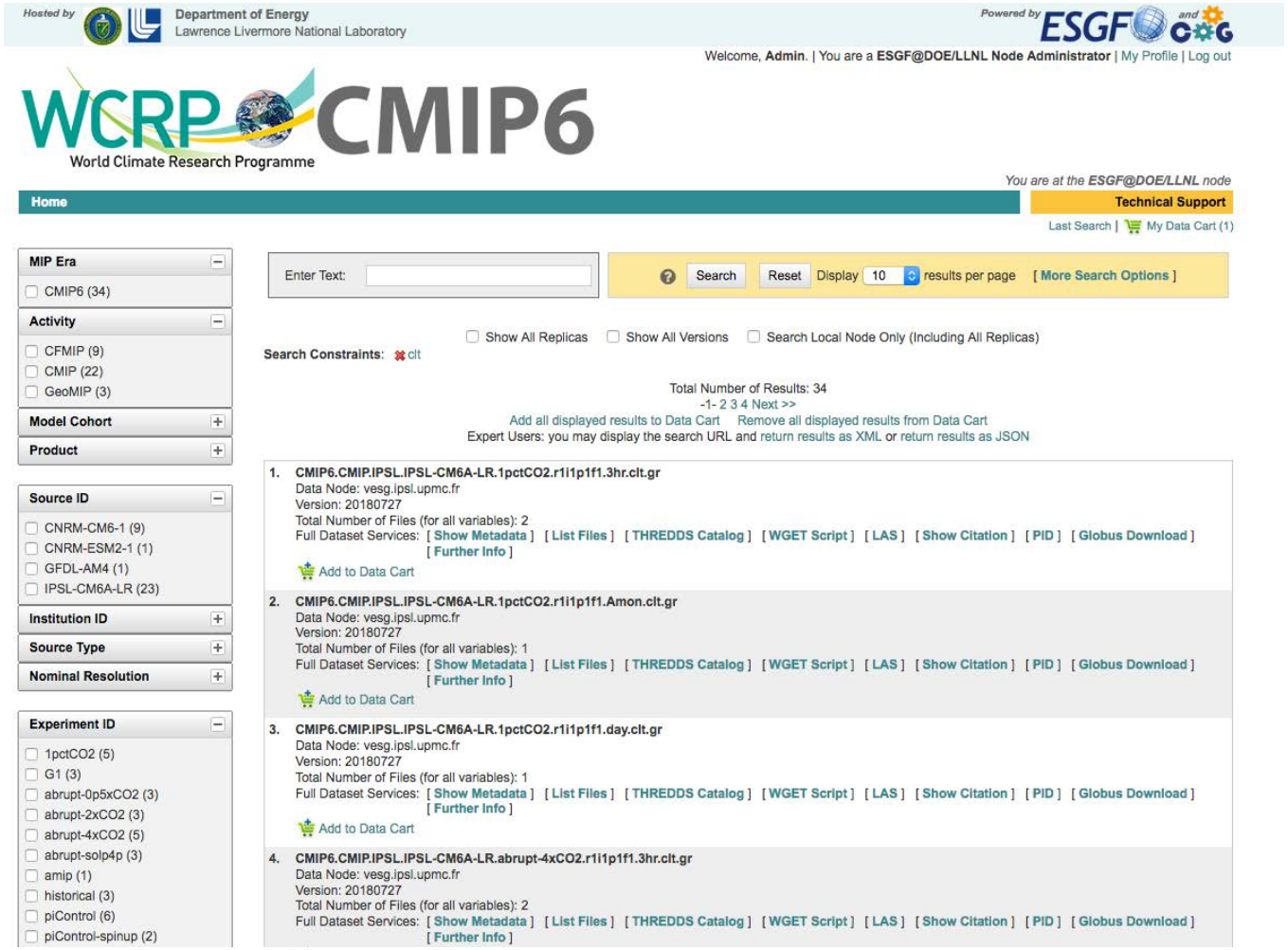
Challenges and Solutions

▪ Distributed Index	→Replicated Solr shards
▪ Trust of distributed identity providers	→Centralized truststore; infrequent dissemination
▪ Data Replication among sites	→Synda software
▪ Bulk data movement	→Globus/Gridftp
▪ Software stack installation	→Yesterday: bash Today: Ansible

Metadata Standards / Curation

- Needed for success of ESGF
- Not managed by ESGF - project specific data preparation steps
- ESGF integrates tools from projects for standards: internal and external metadata checks
- ESGF becomes enforcement mechanism – prevent publishing erroneous data
- CMIP5 came close; CMIP6 universal adoption

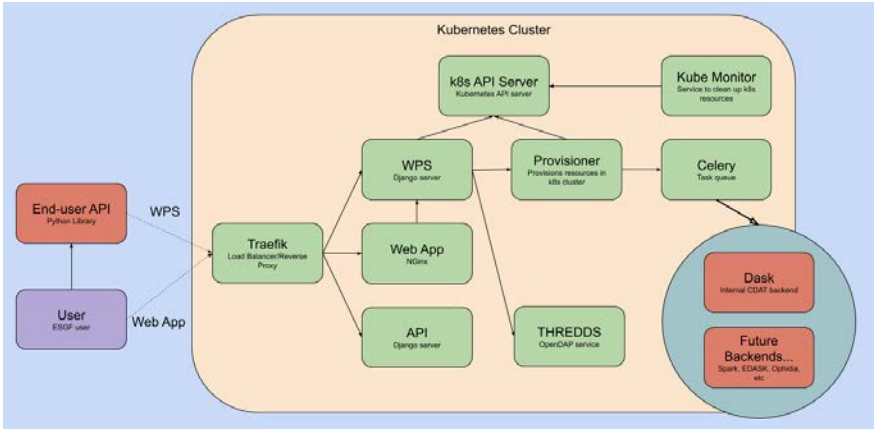
ESGF Web Interface



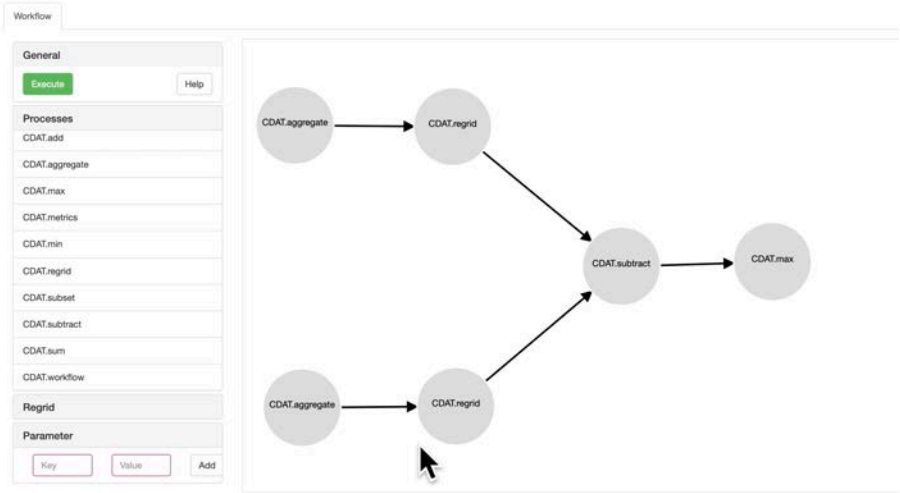
- Faceted search
- Data cart
- Links for metadata and services

ESGF Compute Services

- Next-generation data analyses requires moving analytics close to data archive nodes and enabling server-side computation.
- Create ESGF analytics capabilities by exposing compute resources through well defined interfaces
- Allow ESGF users to download the outputs of analysis rather than huge data sets
- Generic WPS-base API to communicate with server
  - Easy to add new technologies
  - Easy to adapt to any hardware
- API fits multiple backends, eg. DASK or Spark
- The ESGF Compute Cluster hosts LLNL's compute services that provide data sub-setting, add/subtract, and regridding to reduce the amount of overall data transfers.
- Metrics are collected to monitor usage.



LLNL Compute Server Software Architecture



LLNL User Friendly Interface to Compute Services

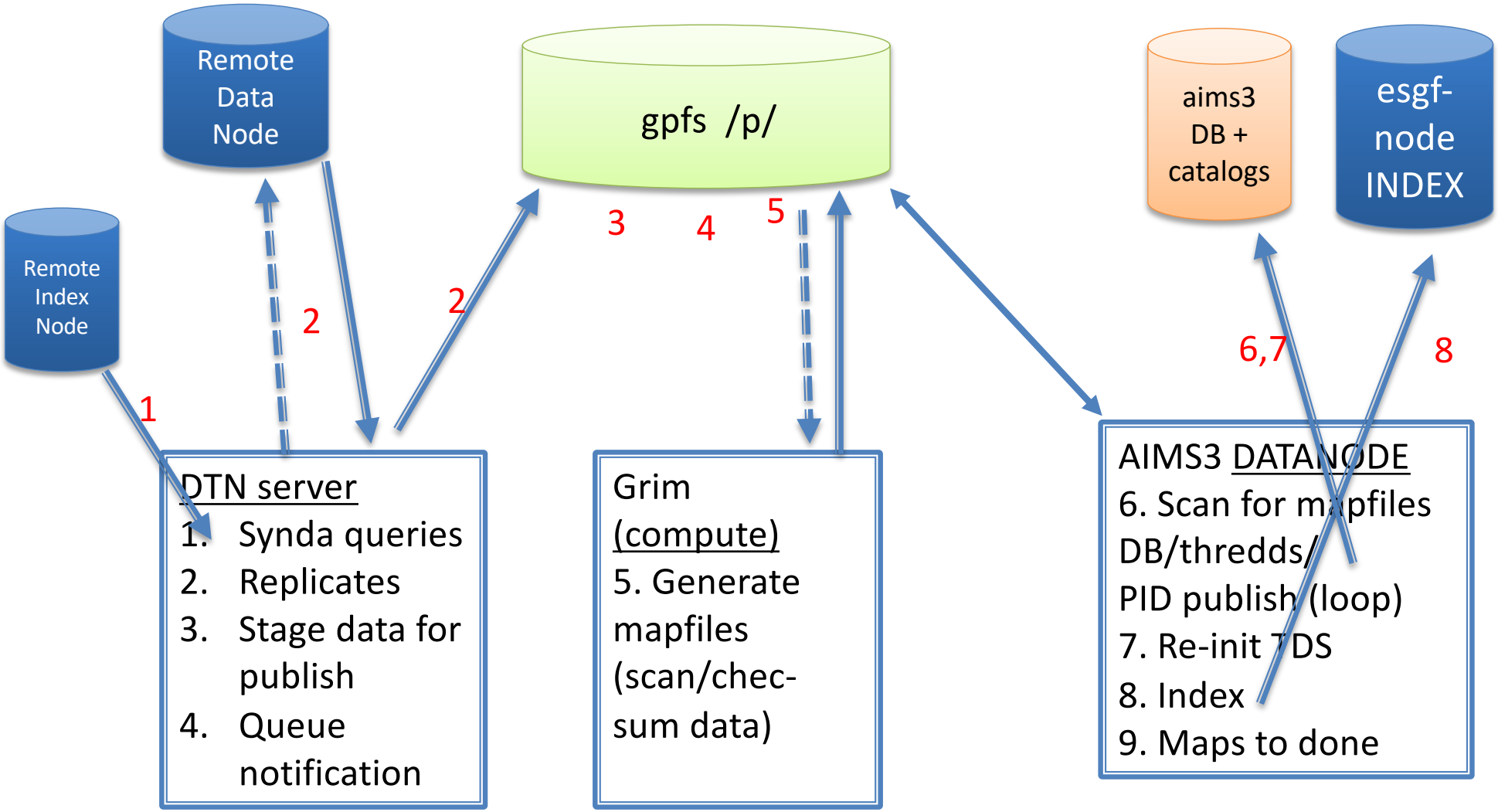
Lessons learned

- ESGF infrastructure is continually required to improve and adapt
- ESGF must continue to rely on careful integration of already proven technologies and applications that have been developed by teams over the course of many years (e.g., Solr, TDS, UV-CDAT, OPeNDAP, etc.)
- Promote participation and involvement by a large community of stakeholders, managers, engineers, through an open source meritocracy based system (not dissimilar to the principles promoted by the Apache Software Foundation, for example)
- Establish a governance model from the very beginning, in order to represent the interests of all stakeholders, prioritize requirements, and guide the overall system development
- Avoid single points-of-failure in the engineering workforce, hardware, software, etc.
- Large infrastructures like ESGF should consider scalability as one of its major requirements (e.g., data discovery, movement, processing, etc. testing should be scaled to 10 to 100 times the current amount of data)
- Funding is always a struggle (US and EU agencies tend to fund innovative research and new ideas and less prone to support ongoing successful projects such as ESGF.

Data Replication at LLNL

- LLNL is major replica site for CMIP data
- 500K+ CMIP6 datasets already published
- Millions more expected; we have to keep up with publication
- Replication happens using Synda
- We have publication tools to publish but no drop-in software exists to handle the volume
- Solution: ”piecemeal” automation
  - Started with several manual steps, automated as time permitted

Replica Workflow



Issues Affecting Workflow

- File systems un-mounted - can't scan data / load mapfiles
- Script bug – fail to move list to done queue, re-issues
- Forget to renew certificate -> solved with auto-renewal
- Performance inconsistencies
  - FS Bandwidth
  - Variability of mapfile generation – slow publication
    - Due to low bandwidth
    - Large data size
    - Small file count == poor parallelism
    - Solution - "prime the pump"
- New model value in data not found in config – halts mapfile generation
- THREDDS (catalog service) becomes unavailable
- Delayed response times for index publish due to heavy external request volume

Rethinking the Architecture

- After 8-10 years, time to revisit requirements
  - Example: control API access to prevent accidental Denial-of-Service
- Consider technologies previously unavailable
  - Mature NoSQL databases
  - Containers
  - Cloud
  - Proliferation of Deep Learning (for compute?)

Conclusions and Further Info

- ESGF addresses the distributed data problem using a federated approach to management
- LLNL plays a lead role in development and data replica management
- Community in process of rethinking architecture
- Website: https://esgf.llnl.gov
- Software stack: https://esgf.github.io/esgf-ansible
- Slack: https://esgf-chat.slack.com