Climate Change Data Management & Distributed Resources

National Academy of Sciences Workshop
Multiple Data Sources

Dean N. Williams
Project Lead for the Analytics and Informatics Management Systems

February 25, 2016

LLNL-PRES-671937
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

Overview

Motivation

Requirements

Approach

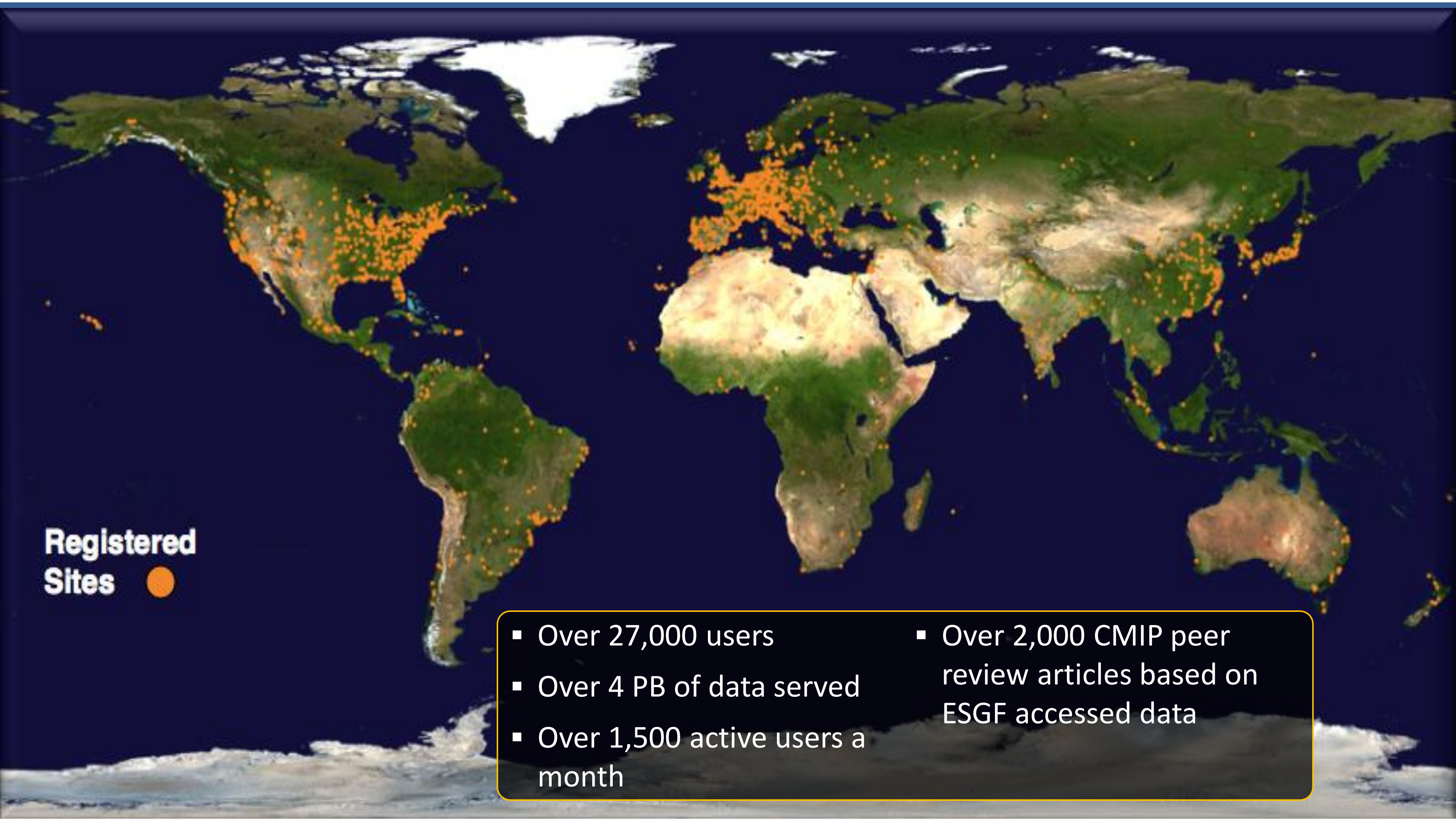
Architecture

Motivating Use Cases

Outreach and Partnerships

ESGF is the world's leading source for climate modeling data

Making data a community resource, accessible worldwide



- Over 27,000 users
- Over 4 PB of data served
- Over 1,500 active users a month
- Over 2,000 CMIP peer review articles based on ESGF accessed data

Automating infrastructure for archiving and comparing simulation results, diagnostics, and validations

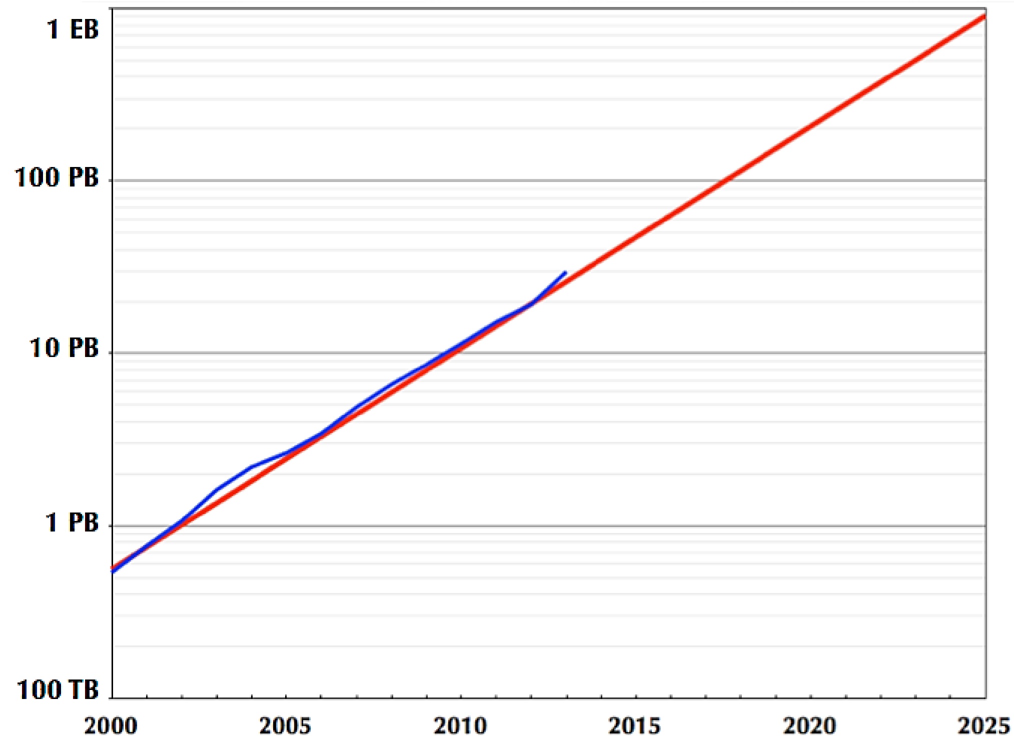
- Leading the creation of a flexible, extensible infrastructure for future national and international climate efforts
- Exploiting broad data sharing for scientific breakthroughs
- Projects represent state of the art in several disciplines
  - Earth System Grid Federation (ESGF)
  - Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT)
    - Climate Compute Working Group (C^2WG)
  - International Climate Network Working Group (ICNWG)
- Empowering advances by integrating our high-quality data streams for a virtual laboratory

ESGF is leading U.S. government agencies (i.e., DOE, NASA, NOAA) and the climate community to integrate all existing and future data holdings into a seamless and unified environment.

Climate "big data" characteristics and challenges

- Petabytes of data distributed around the world
- Requires combining model output with various kinds of observations and reanalysis
- Data and algorithms must be tracked and validated

Community Earth System Model (CESM) archive volume image courtesy of Gary Strand

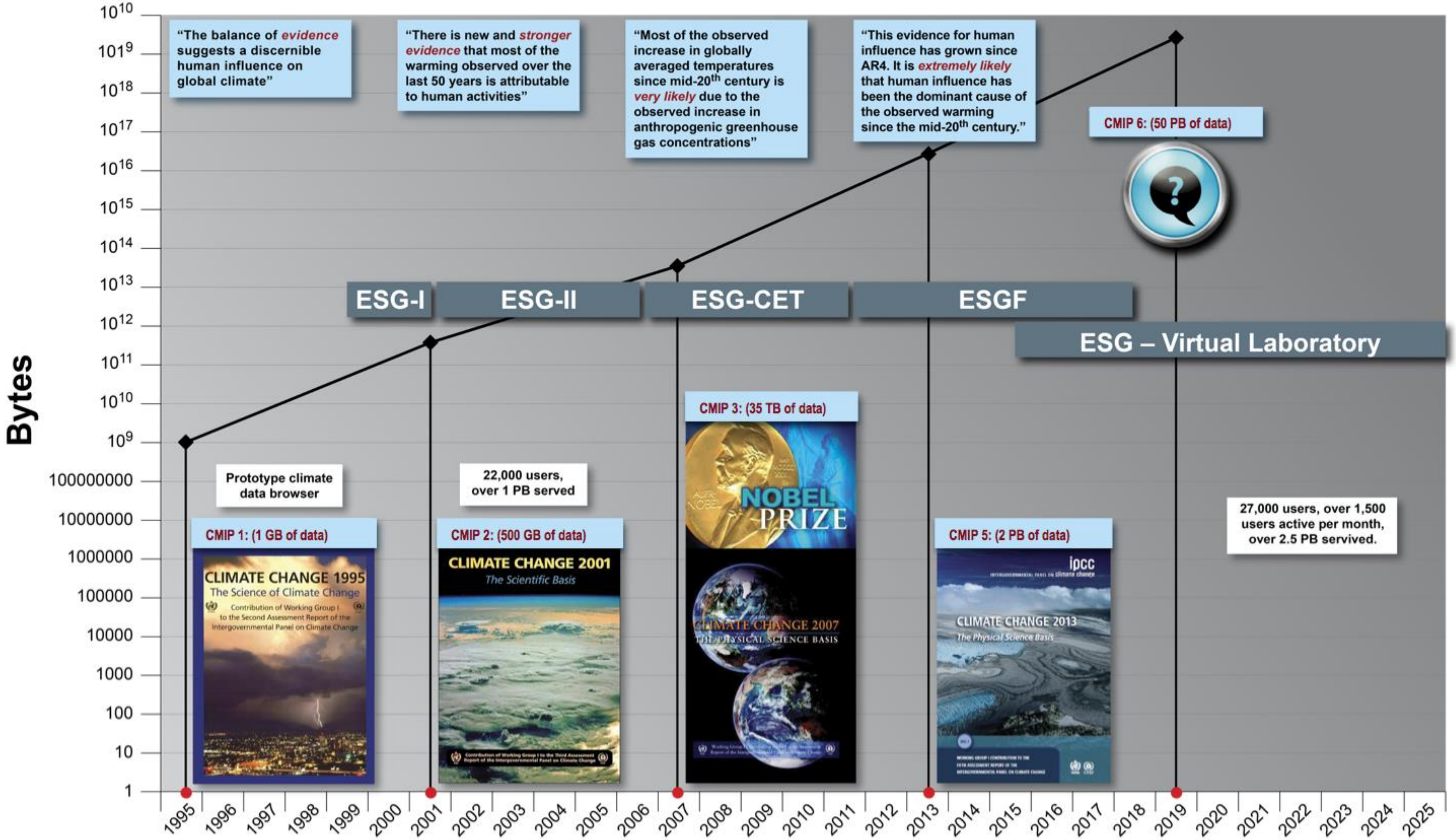


We need new methods to deal with rapid data growth.

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 62 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).

CMIP and ESGF history: scientific challenges and motivation use case



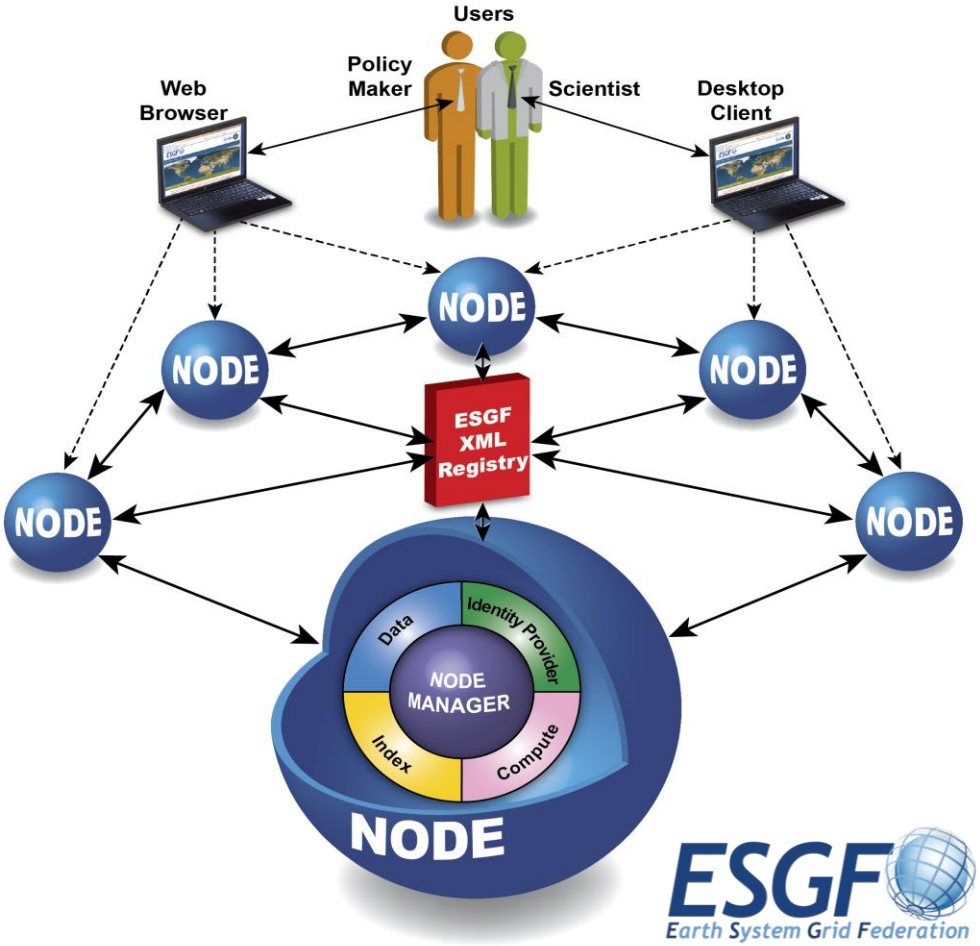
ESGF[1] is a coordinated multiagency, international collaboration of institutions that continually develop, deploy, and maintain software needed to facilitate and empower the study of climate

. . .

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

Federated distributed data archival and retrieval system

- Federated peer-to-peer architecture
- Support discipline-specific portals
- Support browser-based and direct client access
- Single sign-on
- Automated script and GUI-based publication tools
- Full support for data aggregations
- User notification service



ESGF Earth System Grid Federation

Enabling climate research in a data-rich environment

Flexible approach

Our approach generalizes the current operational infrastructure used by the ESGF into a template architecture that hides each implementation layer behind a well-defined Application Programming Interface (API), so that different communities may decide to adopt or swap any single part. The deployable software stack will include the following modules:

- **Publishing services** (reference implementation based on ESGF publishing software);
- **Search services** (reference implementation based on Solr-Cloud and ESGF Search API);
- **Transfer services** (reference implementation based on Globus/GridFTP);
- **Computation services** (reference implementation based on Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT) and ESGF Computing API);
- **Resource monitoring and allocation** (to be developed from scratch);
- **Security services** (reference implementation based on open standards and ESGF security infrastructure);
- **User interface** (reference implementation based on CoG web knowledge environment); and
- **Exploration services** (remote analysis and visualization based on streaming, multi-resolution data)

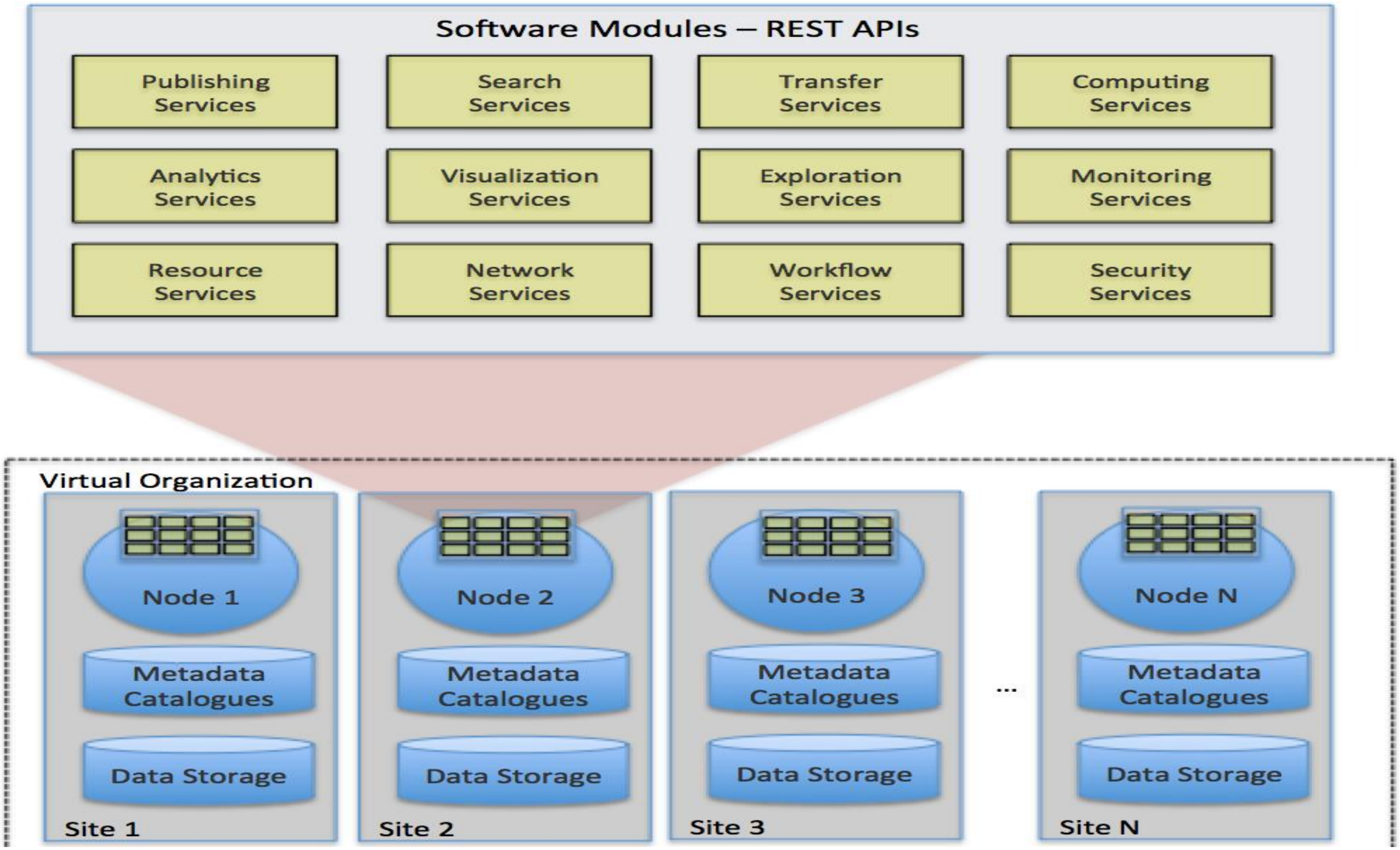
Architecture requirements

The two basic principles behind the architecture design are "**modularity**" and "**abstraction**". Modularity means that architecture will be structured not as a single monolithic system, but rather as a composition of interacting software services, which are packaged and can be installed individually and independently

The functionality of each service will be abstracted in a well-defined API, so that each service can be easily invoked by other services and clients without worrying about the underlying implementation details. All service APIs will be defined to conform to the Representational State Transfer (REST) web service paradigm, which will allow simple invocation by standard web-enabled clients

In general, software systems that are both modular and abstract have an intrinsic longer longevity, because each service can be evolved or replaced individually, without affecting the backward compatibility with other parts of the systems, or its clients.

Architecture



Achieving community scientific goals requires additional storage and computing resources, along with a common virtual computational environment that conforms to established standards across the federation.

Architecture description

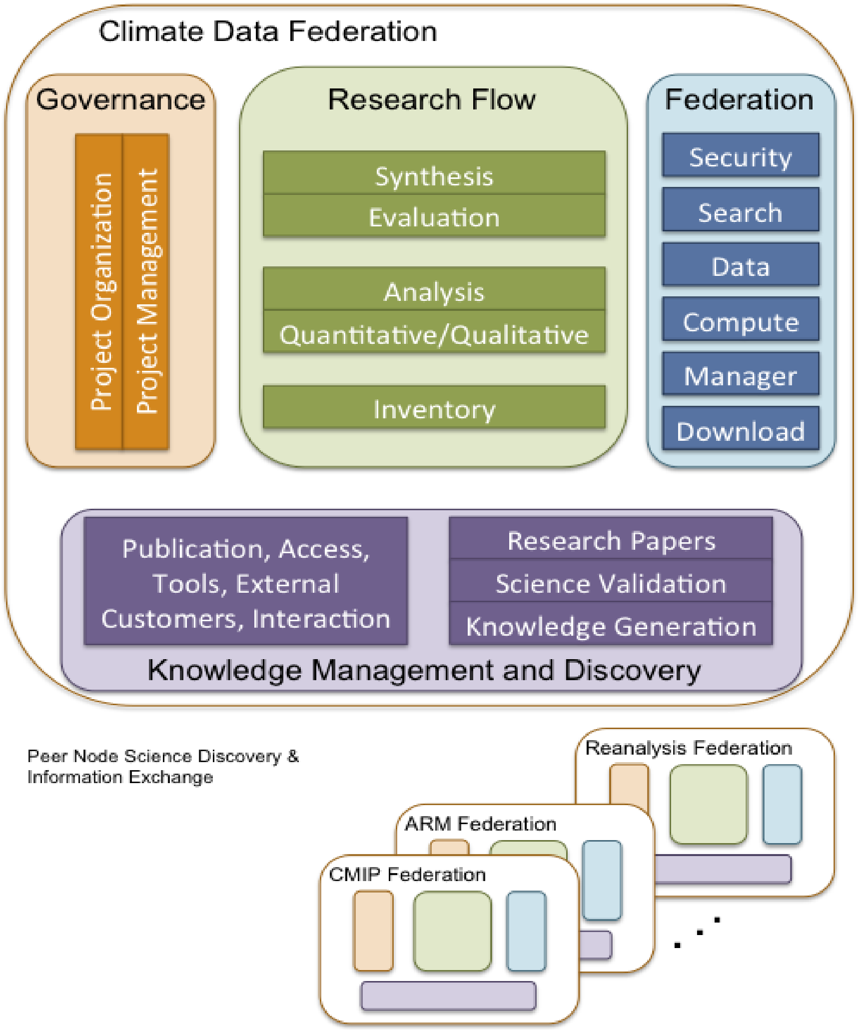
- **Distributed**: the architecture must support a system composed of geographically distributed nodes, each with its own set of data resources, metadata catalogs, and software services. Protocols and services must unite all nodes in a single federation, so that a client can discover, search, download and execute resources independent of their physical location, as if they were held and run on a single server.
- **Dynamic**: the architecture must be designed to support a highly dynamic system, which will expose with minimum delay all data and services that are available at that time throughout the federation. For example, data collections produced by a model run, data streams originating from the real-time processing of a field instrument, or new derived products produced by a scientist running their processing code, could all be published into the system and immediately returned as results of federation-wide searches. Additionally, the system must continually provide an up-to-date report on the state of its components and must be able to automatically direct client requests where resources are available.
- **Scalable**: the architecture must be able to scale to the "Big Data" volumes that are expected in several scientific fields (climate, astronomy, genomics, etc.) in the next 5-10 years. Scalability must be achieved through a two-fold approach. First, each service must be implemented through a high-performance technology that is inherently able to handle large volumes of data, in a distributed environment. Second, the modular architecture must allow each service to be instantiate multiple times.

Architecture description

- **Resiliency and Fault Tolerance**: The architecture must be designed to include redundant components for all critical services (such as search, authentication, authorization, data download and visualization), and to execute automatic failover in case any component becomes unavailable. When backup services are not available, it must ensure that it produces meaningful error responses to human and machine clients.
- **Secure**: The architecture must support a distributed and federated security model, whereby each node will maintain complete control over the policies for accessing its local data and computational resources, while federation-wide authentication and authorization services are responsible for enforcing these policies homogenously through the system. Single-sign-on and federated access control must allow users to register and authenticate only once, and then propagate their identity.

Sharing infrastructure burden through common and consistent standardization and peer-to-peer services

- NetCDF climate and forecast (CF) metadata convention
- Climate Model Output Rewrite 2 (CMOR-2)
- Regridders: Earth System Modeling Framework (ESMF), Spherical Coordinated Remapping and Interpolation Package (SCRIP), GRIDSPEC
- Publishing
- Search and discovery
- Replication and transport
- Common information model
- Quality control
- Website and web portal
- Security
- Product services



ESGF sets networking best practices into place to effectively transport tens of petabytes of climate data

Immediate goal: 4 Gbps (1 PB/month) of sustained disk-to-disk data transfer between ESGF primary data centers

Stretch goal: 16 Gbps (1 PB/week) of sustained disk-to-disk data transfer between ESGF primary data centers

CMIP6 data lifecycle

Most of the work is currently spent in coordination of CMIP6 output and ESGF data publication:

- Lists of requested variables
- Controlled vocabularies
- Global attributes
- NetCDF file requirements for ESGF publication
- Idea: Model Output Preparation Package

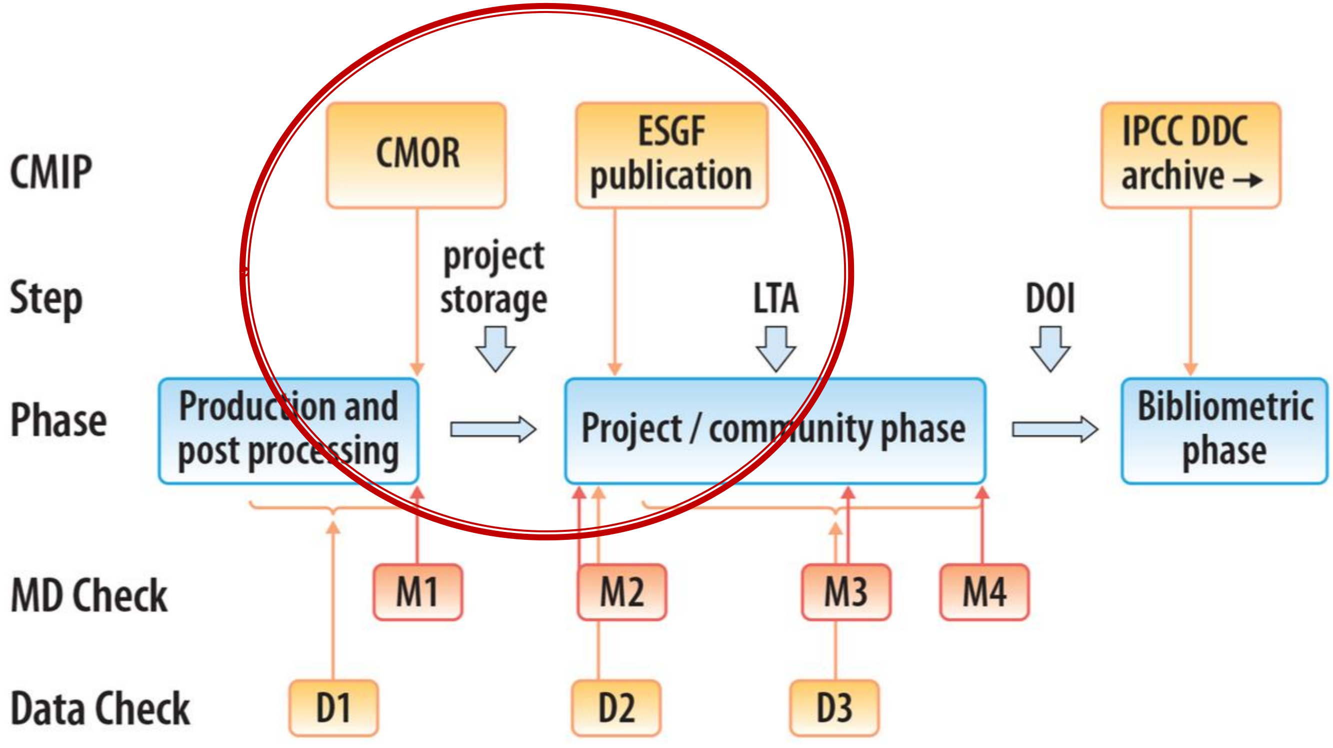
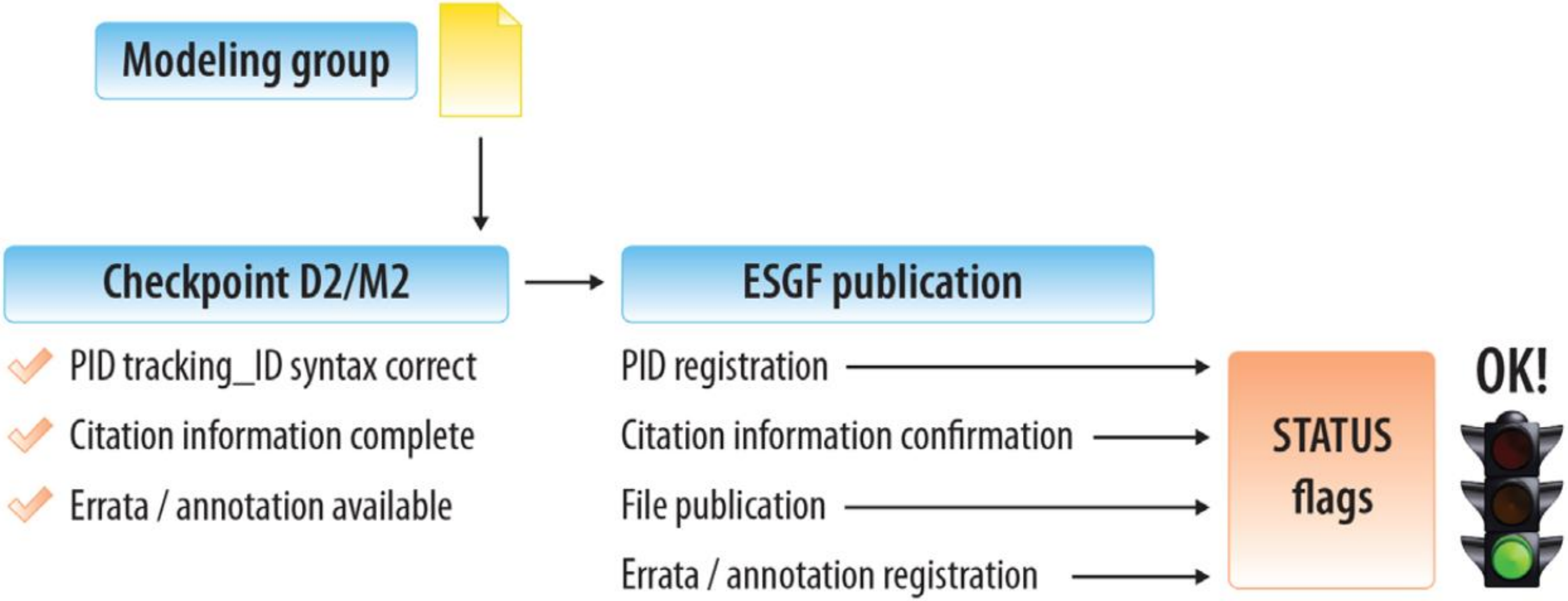


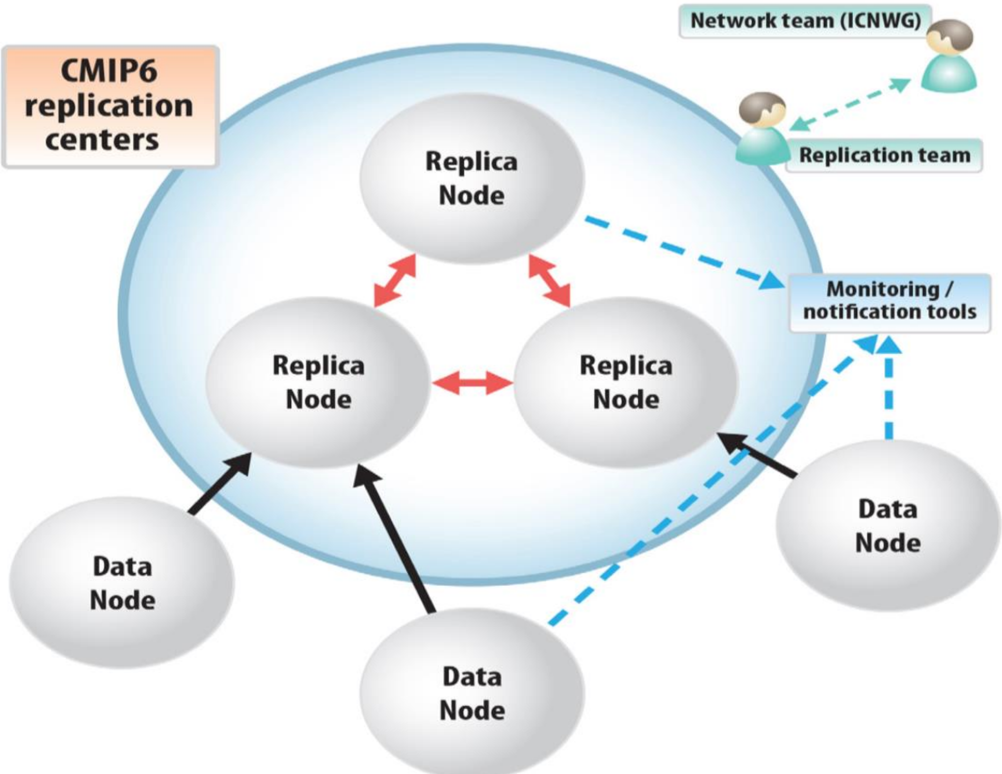
Diagram from QA WIP position paper. (D1,D2...Data Checks, M1,M2... Metadata Checks, QC of software not represented)

CMIP6 data lifecycle



Schematic CMIP6 Handle PID and early citation data publication service. The services for PID and citations information have to be completed by the service on errata and annotations as the third new service in addition to the classical ESGF file publication. Details on these new ESGF services for CMIP6 can be obtained from the three corresponding WIP position papers.

CMIP6 data lifecycle



CMIP6 replication from data nodes to replica centers and between replica centers coordinated by a CMIP6 replication team. More details can be obtained from the corresponding WIP position paper.

Data replication is not only a technical problem with respect to network bandwidth but also a management problem with respect to data consistency and replication strategy. Probably not all CMIP6 core data can be stored in one location.

ESGF strategic roadmap of proposed components in the integrated data ecosystem

ESGF data ecosystem roadmap

Current: Interoperable & Distributed Databases — **Distributed Data Ecosystem** — **Virtual Laboratory**

Planned ESGF Data System Evolution

2016

Federated databases

- Peer-to-peer data archive
- Synchronized federated metadata and data
- Easy "fire and forget" large-scale data transfer
- Quick-look server-side analysis
- Virtual organization Management
- Content management, conventions, and standard web portal
- Secure access control

Early 2018

Ecosystem data sharing

- Ensemble and parallel server-side analysis, diagnostics, metrics and visualization
- Data quality and citation
- Dynamic resource management
- Full workflow and provenance capture
- User interface portal scripting
- Network transfer speed between data centers 1 PB/day
- Automated publication, replication, and version control

2022

Virtual Laboratory

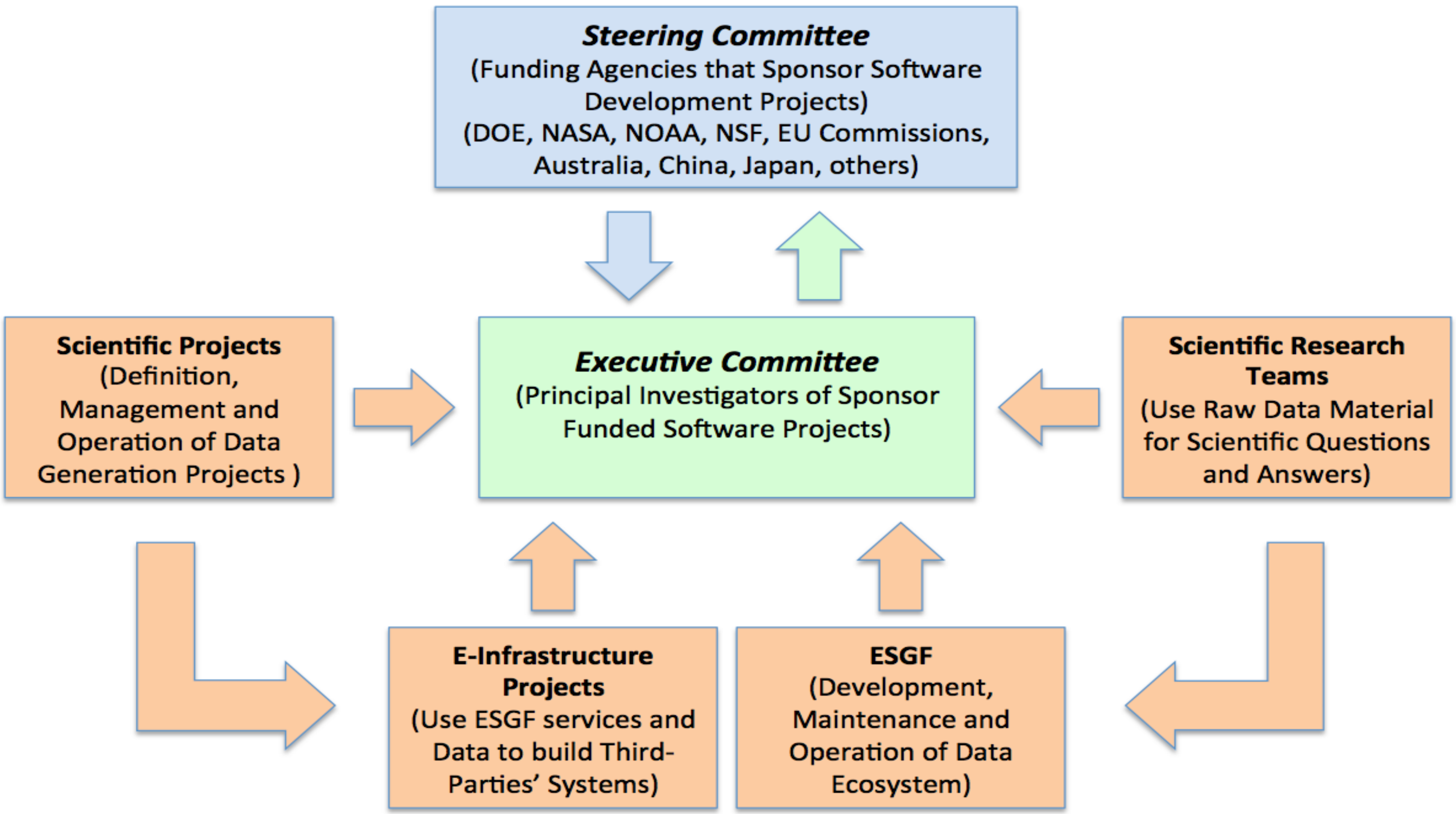
- Uncertainty quantification and derived data products
- Ontology
- Model intercomparison metrics
- Local and remote exploratory data analytics, such as pattern discovery, machine learning
- User support, life-cycle maintenance
- Systemwide analytical modeling
  - Understanding and predicting use

MIPs → **ESGF Data Archive** → satellite, in situ climatology, diagnostics ecosystems

Petabytes (10^{15}) — **Exabytes (10^{18})**

Managing ESGF for success

Governance communication architecture

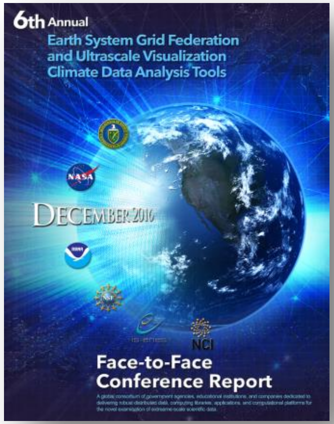


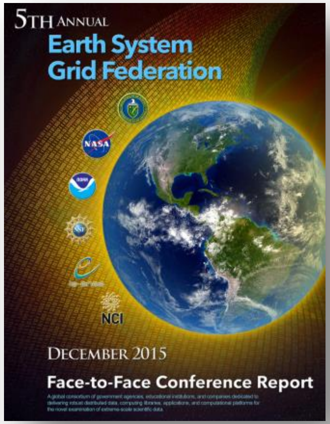
ESGF sub-tasks and task leaders

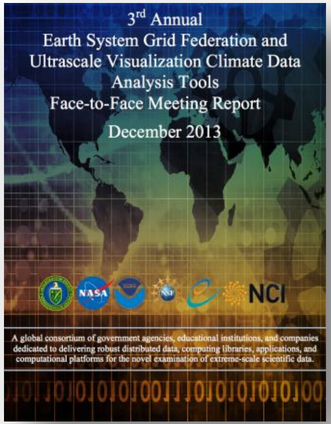
Sub-Task	Task Leads	Description
1. CoG User Interface Working Team	Cecelia DeLuca (NOAA) and Luca Cinquini (NOAA)	Improved ESGF search and data cart management and interface
2. Compute Working Team	Charles Doutriaux (DOE) and Daniel Duffy (NASA)	Developing the capability to enable data analytics within ESGF
3. Dashboard Working Team	Sandro Fiore (IS-ENES)	Statistics related to ESGF user metrics
4. Data Transfer Working Team	Lukasz Lacinski (DOE) and Rachana Ananthakrishnan	ESGF data transfer and enhancement of the web-based download
5. Documentation Working Team	Matthew Harris (DOE) and Sam Fries (DOE)	Document the use of the ESGF software stack
6. Identity Entitlement Access	Philp Kershaw (IS-ENES) and Rachana Ananthakrishnan (DOE)	ESGF X.509 certificate-based authentication and improved interface
7. Installation Working Team	Nicolas Carenton and Prashanth Dwarakanath (IS-ENES)	Installation of the components of the ESGF software stack
8. International Climate Network Working Group	Eli Dart (DOE/ESnet) and Mary Hester (DOE/ESnet)	Increase data transfer rates between the ESGF climate data centers
9. Metadata and Search Working Team	Luca Cinquini (NASA)	ESGF search engine based on Solr5; discoverable search metadata
10. Node Manager Working Team	Sasha Ames (DOE) and Prashanth Dwarakanath (IS-ENES)	Management of ESGF nodes and node communications
11. Provenance Capture Working Team	Bibi Raju (DOE)	ESGF provenance capture for reproducibility and repeatability
12. Publication Working Team	Sasha Ames (DOE) and Rachana Ananthakrishnan	Capability to publish data sets for CMIP and other projects to ESGF
13. Quality Control Working Team	Martina Stockhause (IS-ENES) and Katharina Berger (IS-ENES)	Integration of external information into the ESGF portal
14. Replication Working Team	Stephan Kindermann (IS-ENES) and Tobias Weigel (IS-ENES)	Replication tool for moving data from one ESGF center to another
15. Software Security Working Team	Prashanth Dwarakanath (IS-ENES) and Laura Carriere (NASA)	Security scans to identify vulnerabilities in the ESFF software
16. Tracking / Feedback Notification Working Team	Sasha Ames (DOE)	User and node notification of changed data in the ESGF ecosystem
17. User Support Working Team	Torsten Rathmann (IS-ENES) and Matthew Harris (DOE)	User frequently asked questions regarding ESGF and housed data
18. Versioning Working Team	Stephan Kindermann (IS-ENES) and Tobias Weigel (IS-ENES)	Versioning history of the ESGF published data sets

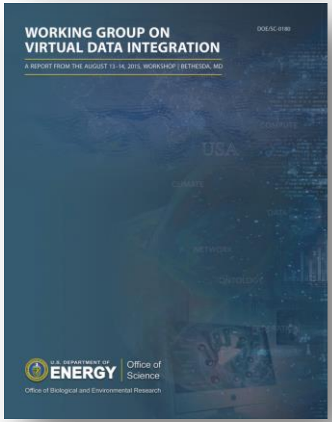
Further elaborations of the sub-tasks are described in the ESGF progress reports, which can be found online: http://esgf.llnl.gov/reports.html

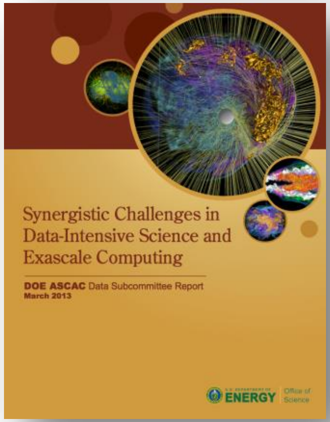
Data workshop and conferences reports: community involvement and outreach

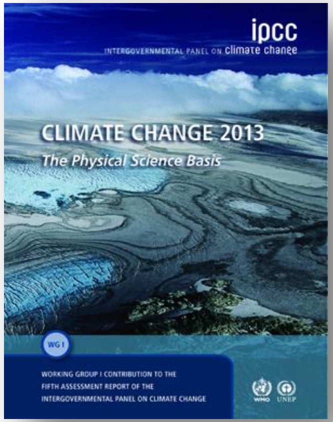


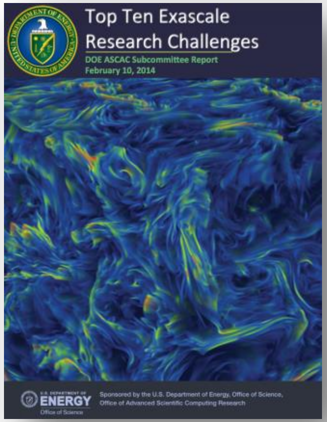












DOE BER CESD workshop and conference reports can be found at: http://esgf.llnl.gov/reports.html and http://science.energy.gov/ber/community-resources/.

Report findings

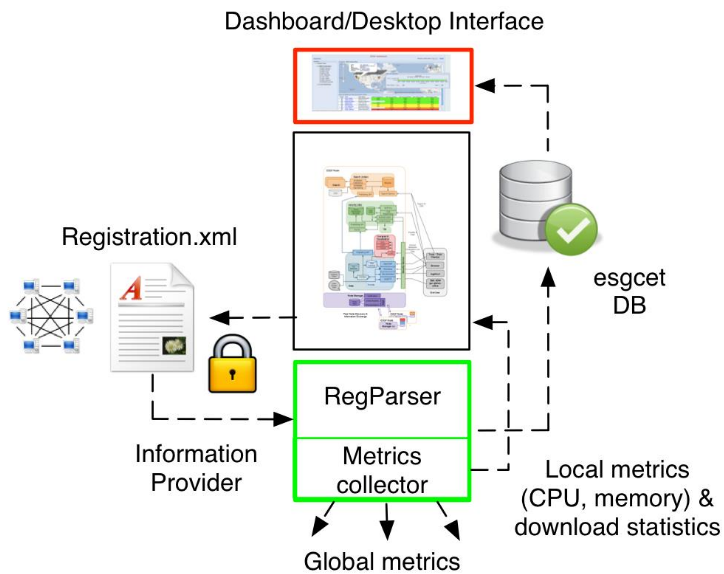
Findings	Description
Data Quality	ESGF data quality persists in the form of provenance, quality control (QC) checks, errata, and data citations. Various components help to improve data quality checks in the ESGF publishing process. EzCMOR (Climate Model Output Rewriter) is one such software package that may be connected to ESGF to enable QC checking before publishing to ESGF.
Data Compression	Data compression is important to ESGF in terms of data storage and transfers. Because of the sheer size of ESGF archives, compressing data for storage or transfer considerably reduces overall costs.
Data Storage	The sheer size of current and expected future archives makes storage a difficult issue to address. If the expected storage for CMIP6 is over 10 petabytes (with estimates as high as 50 PB), then a uniform storage strategy must be put into place among the major CMIP data center sites. This includes the purchasing of storage units and possible archiving of data on tape for long-term data preservation.
Hardware	A cost-benefit analysis is needed for long-term storage. For example, what would it cost to regenerate versus store the data?
Network	ESGF requires the ability to control the timing of data- and network-intensive replication operations for large climate data sets.
Operations	Operational support is needed to sustain the numerous ESGF nodes operated by simulation, observation, and reanalysis projects.
Performance Metrics	Performance metrics must be included as part of ESGF operations. The goal is to have display- able and well-understood performance metrics to track and monitor the overall system and to gather data transfer performance metrics among major CMIP data center sites. Performance operations also must include overall system robustness monitoring and functionality bench- marking to ensure end-user and project satisfaction.
Provenance Capture	Provenance capturing is necessary for reproducing complex analysis processes at various levels of detail in a shared environment.
Server-Side Analysis (and Derived Data Sets)	The size of some data sets makes moving most of the needed data to the end user's home institution infeasible. Data analysis therefore must be performed remotely.
Software Security Scans	The latest software security breach has necessitated an inventory of all software in the ESGF software stack, and ESGF developers have coordinated component development to combine and share information about existing vulnerabilities that may affect secure ESGF operations.
Training and Documentation	Training is important to ensure proper data use and dissemination.
Use Metrics	Use metrics help projects know how the community is using their hardware, software, network, data, and other resources. Metric information such as number of users will serve as base metrics for various data and services within ESGF. Service-specific metrics also should be defined to measure the usage and adoption of specific capabilities and to evaluate their usefulness. Another important metric is identification of the number of software packages provided by other institutes accessible via ESGF.

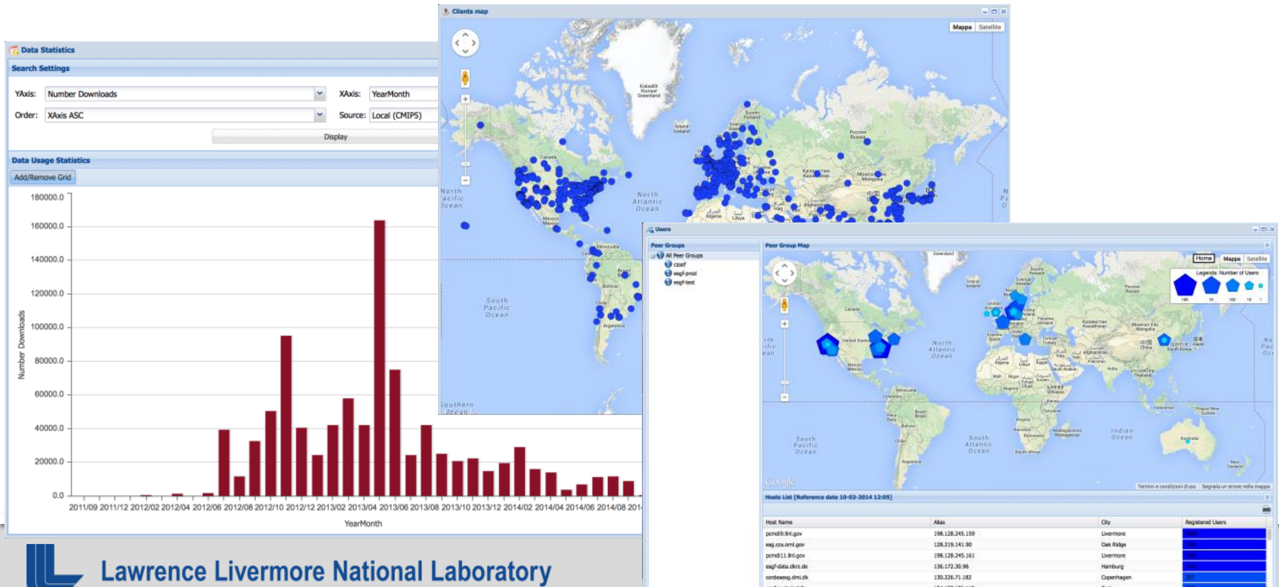
Lessons learned

- ESGF infrastructure is under constant requirements 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
- 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.

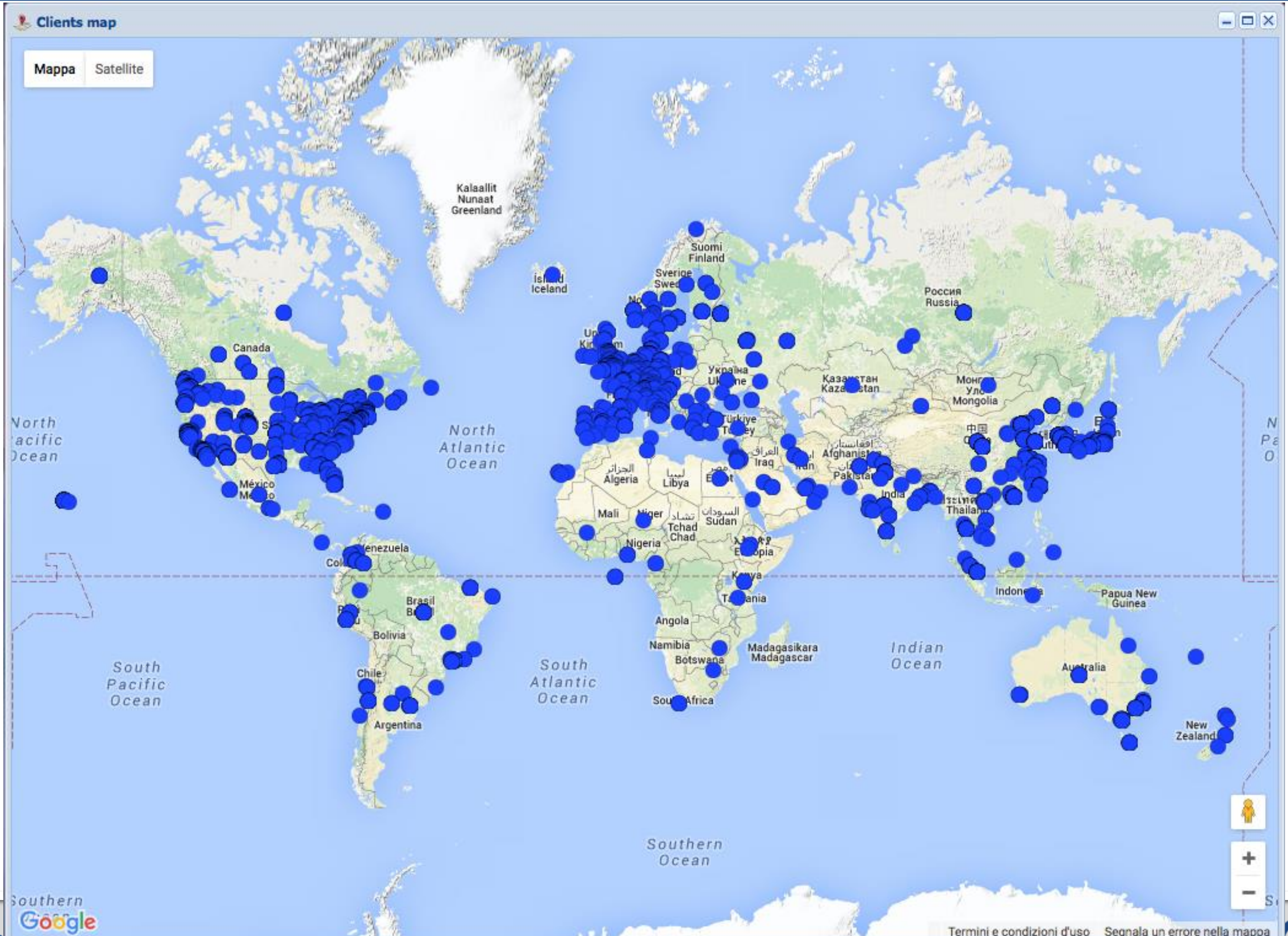
Dashboard Working Team

- ✓ **Development of a monitoring & data usage statistics system**
  - ✓ Coarse grain statistics
    - ✓ Data downloads
    - ✓ Number of users
    - ✓ Number of files
  - ✓ Fine grain statistics
    - ✓ By variable, model, experiment, etc.
    - ✓ CMIP5, CORDEX
    - ✓ EU and Federation-level



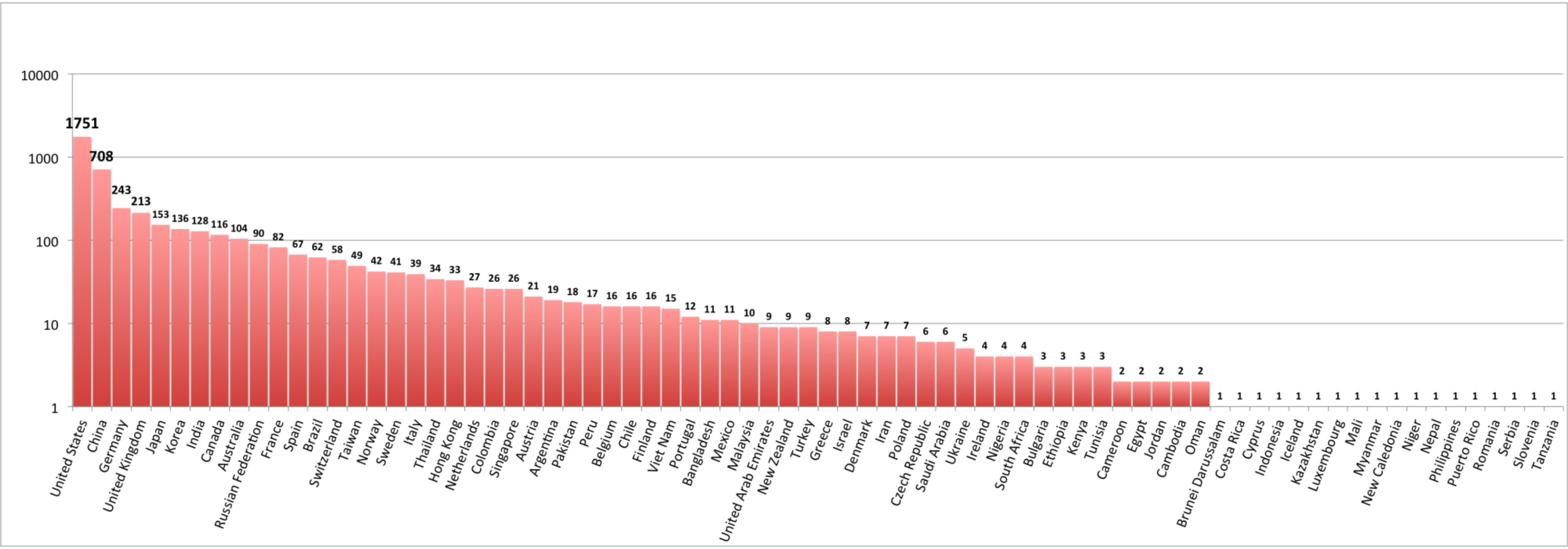


Clients geographic distribution
(LLNL data node view)

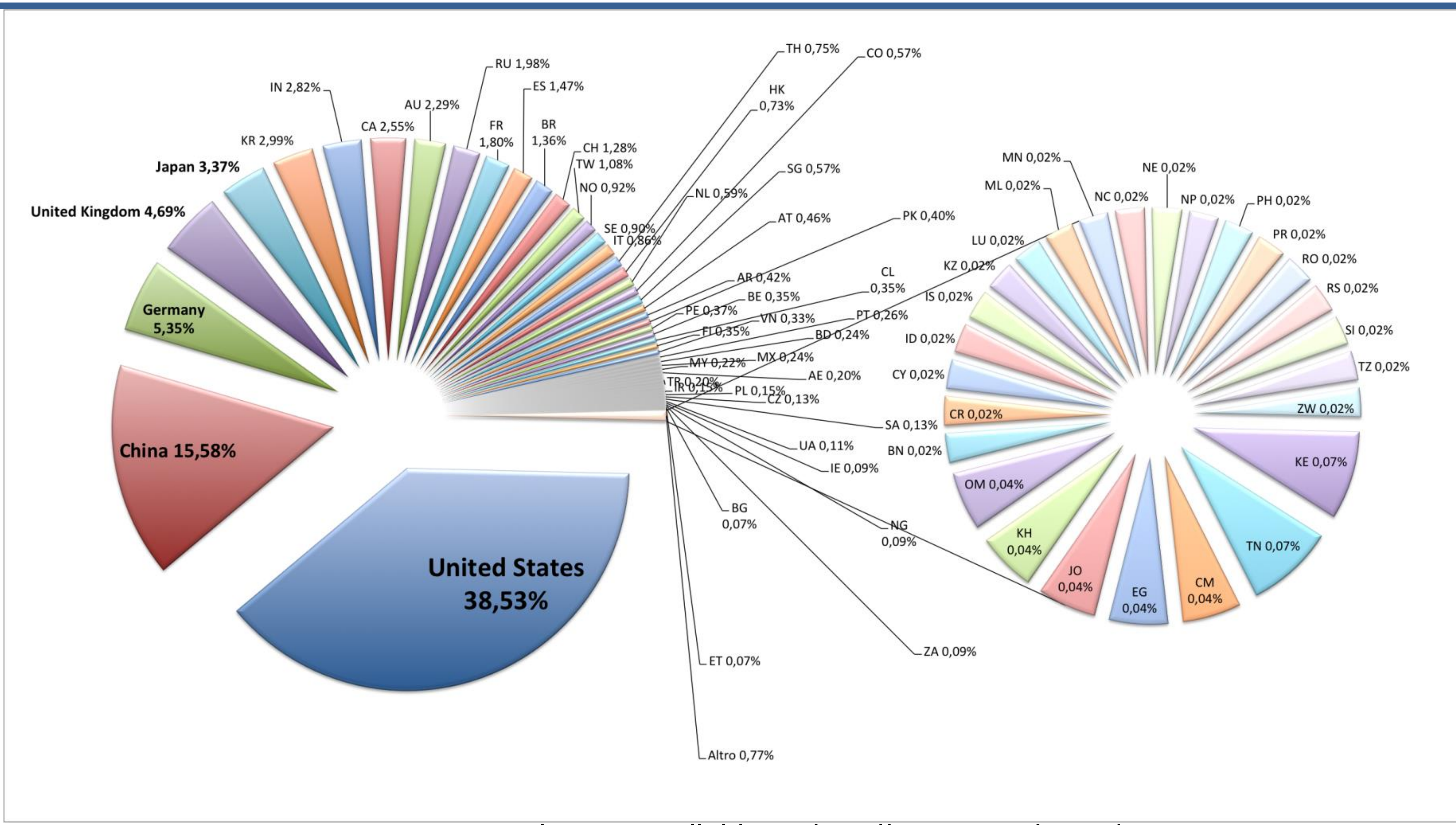


Clients distribution by country
(LLNL data node view)

4650 distinct IPs in total

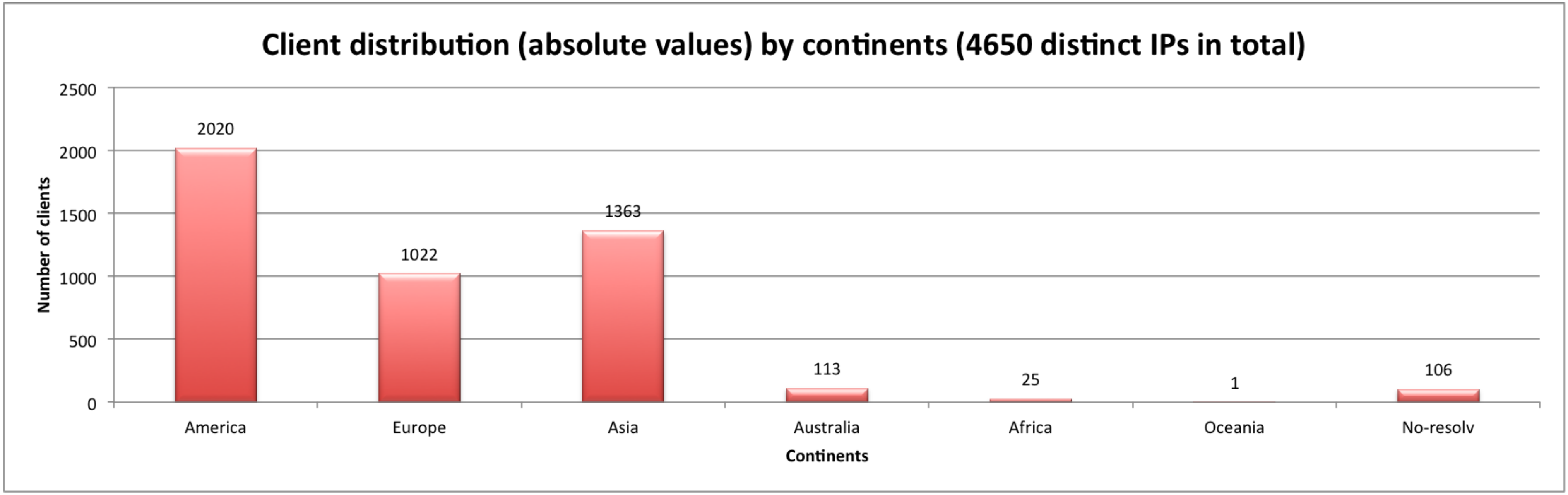


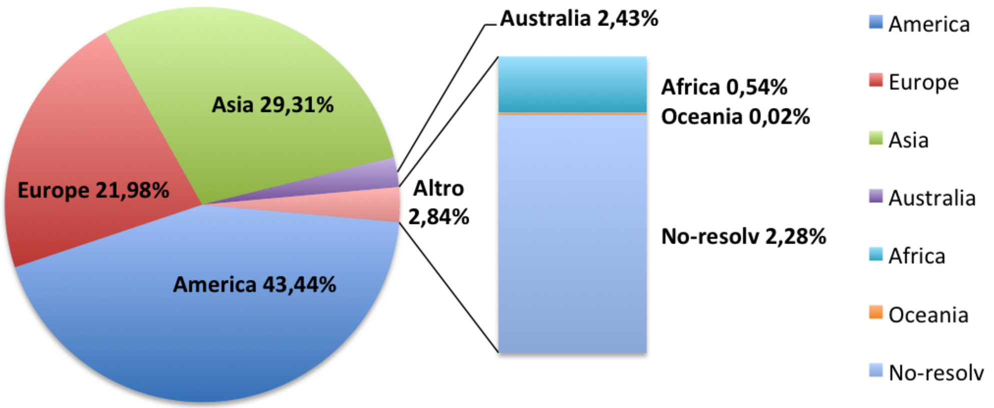
Clients distribution by country (%)
(LLNL data node view)



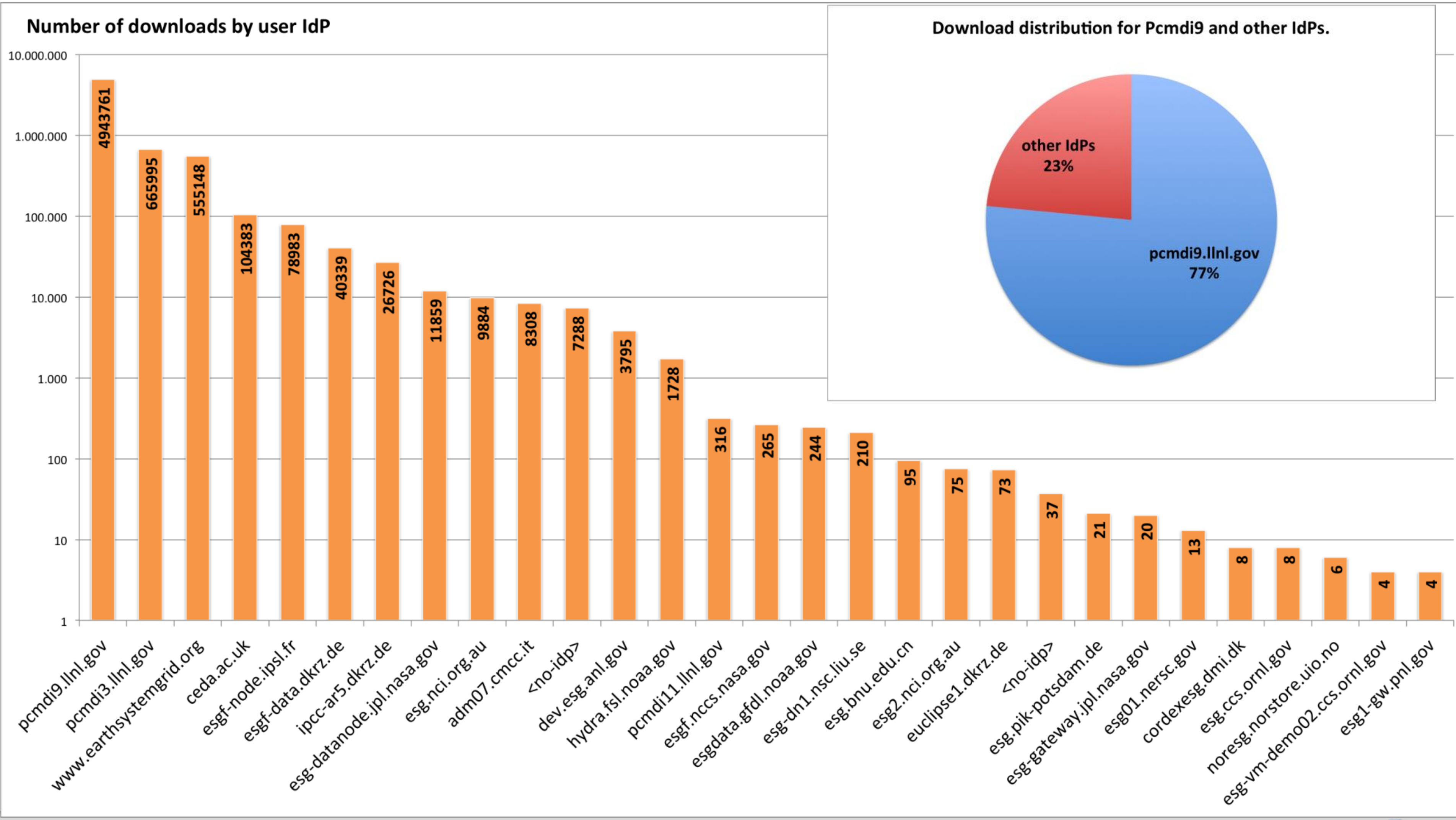
Country codes are available at: http://countrycode.org/

Clients distribution by continent
(LLNL data node view)

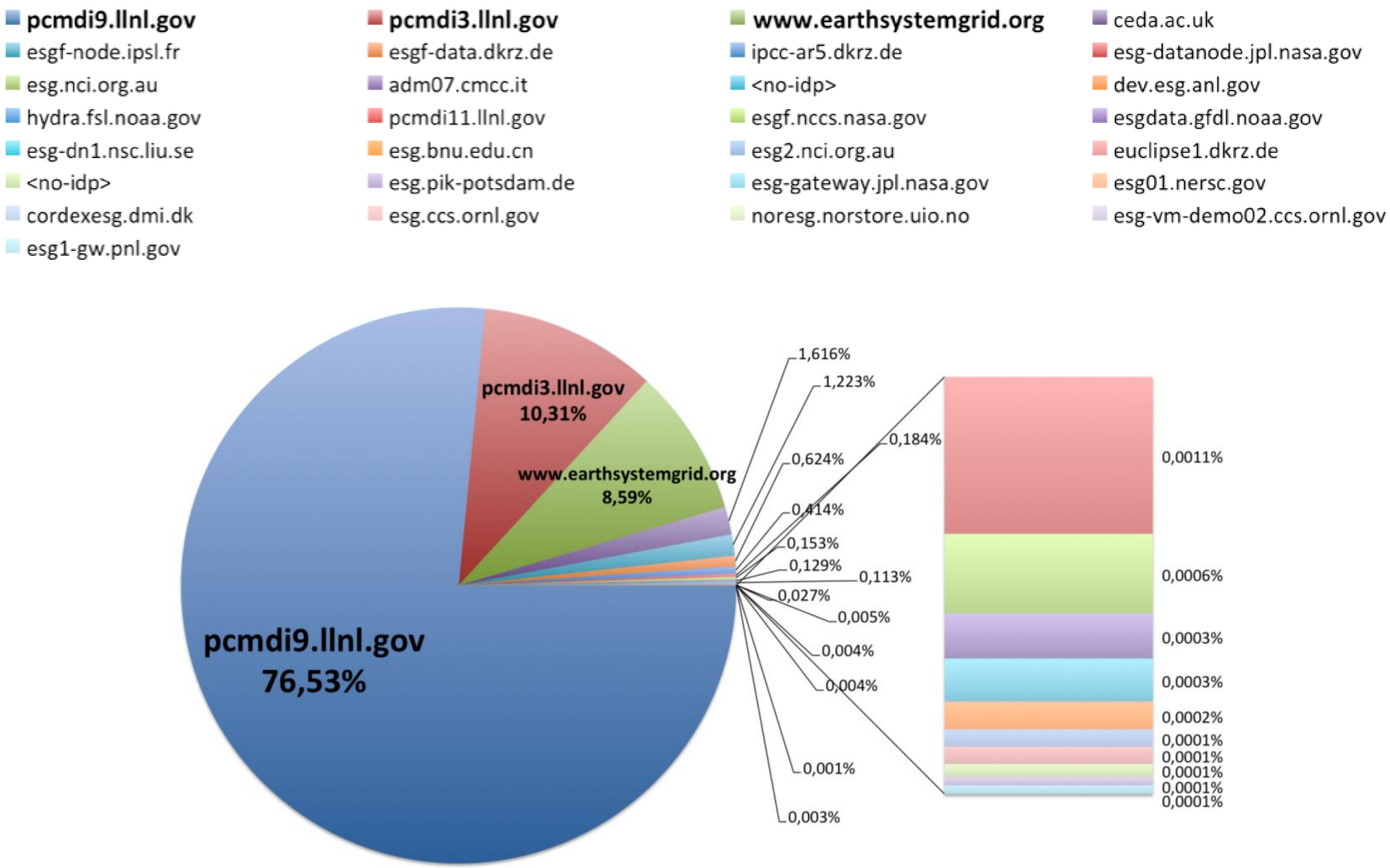




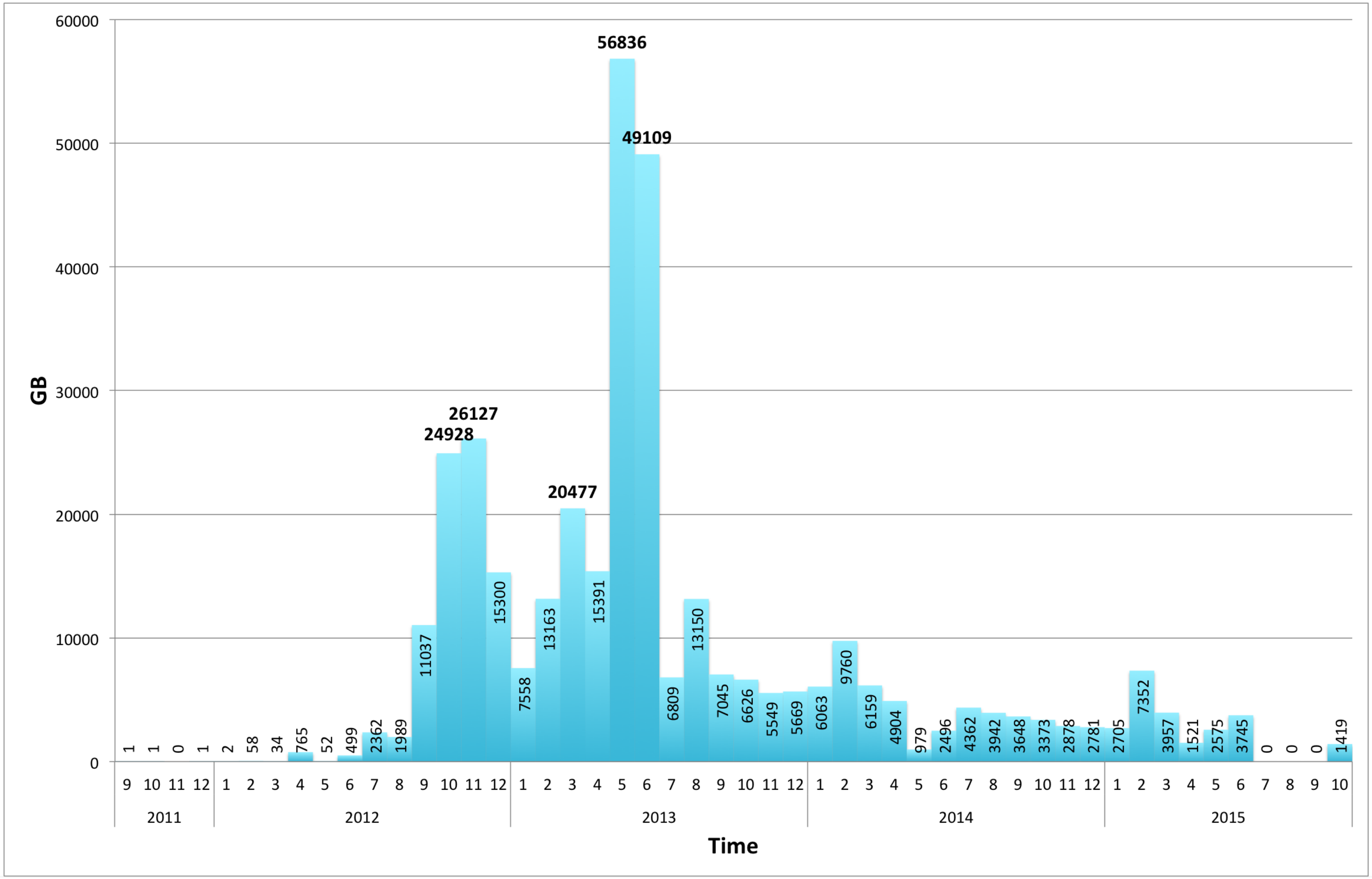
Download distribution by identity provider
(LLNL data node view)



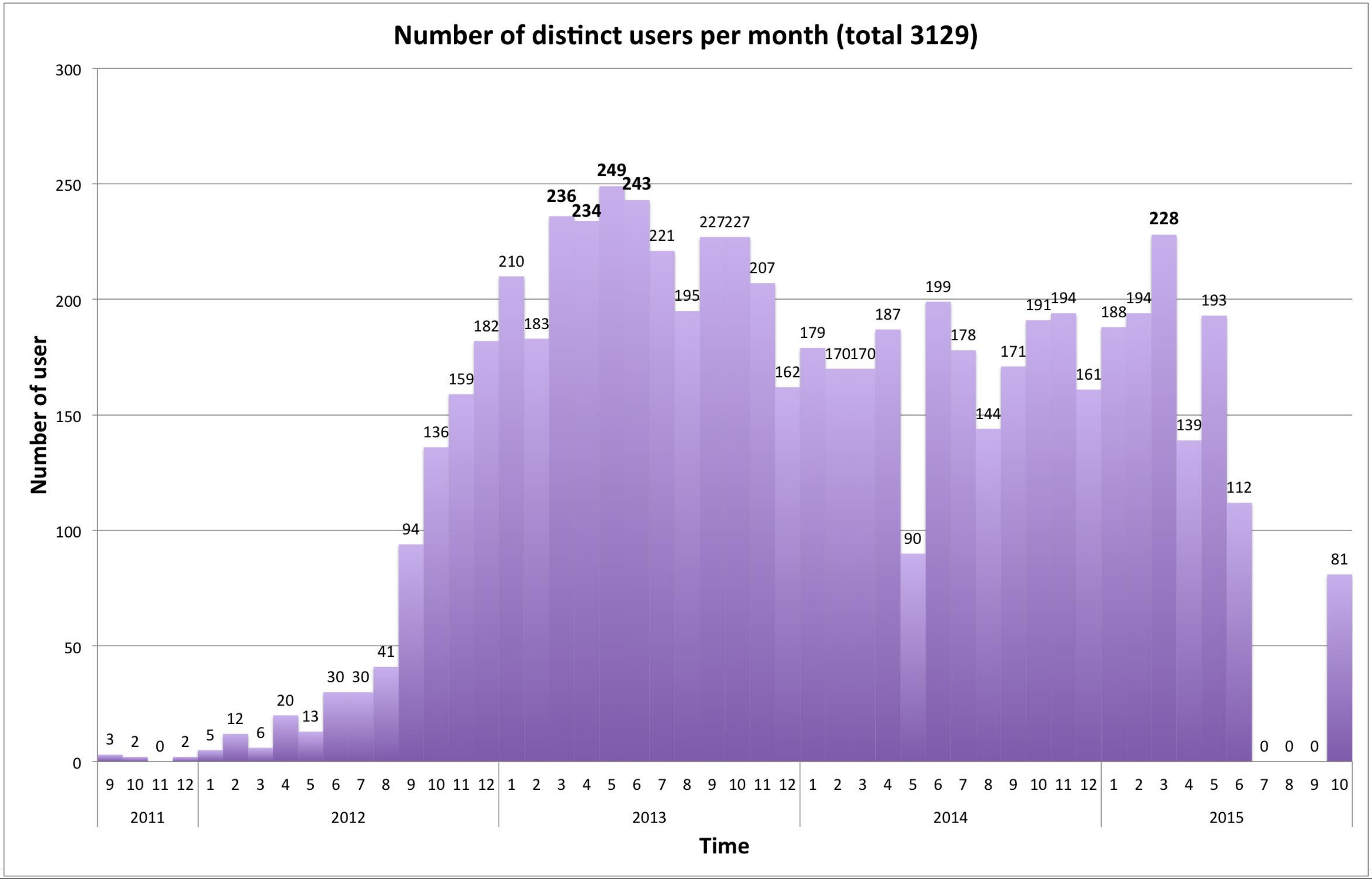
Download distribution (%) by identity provider
(LLNL data node view)



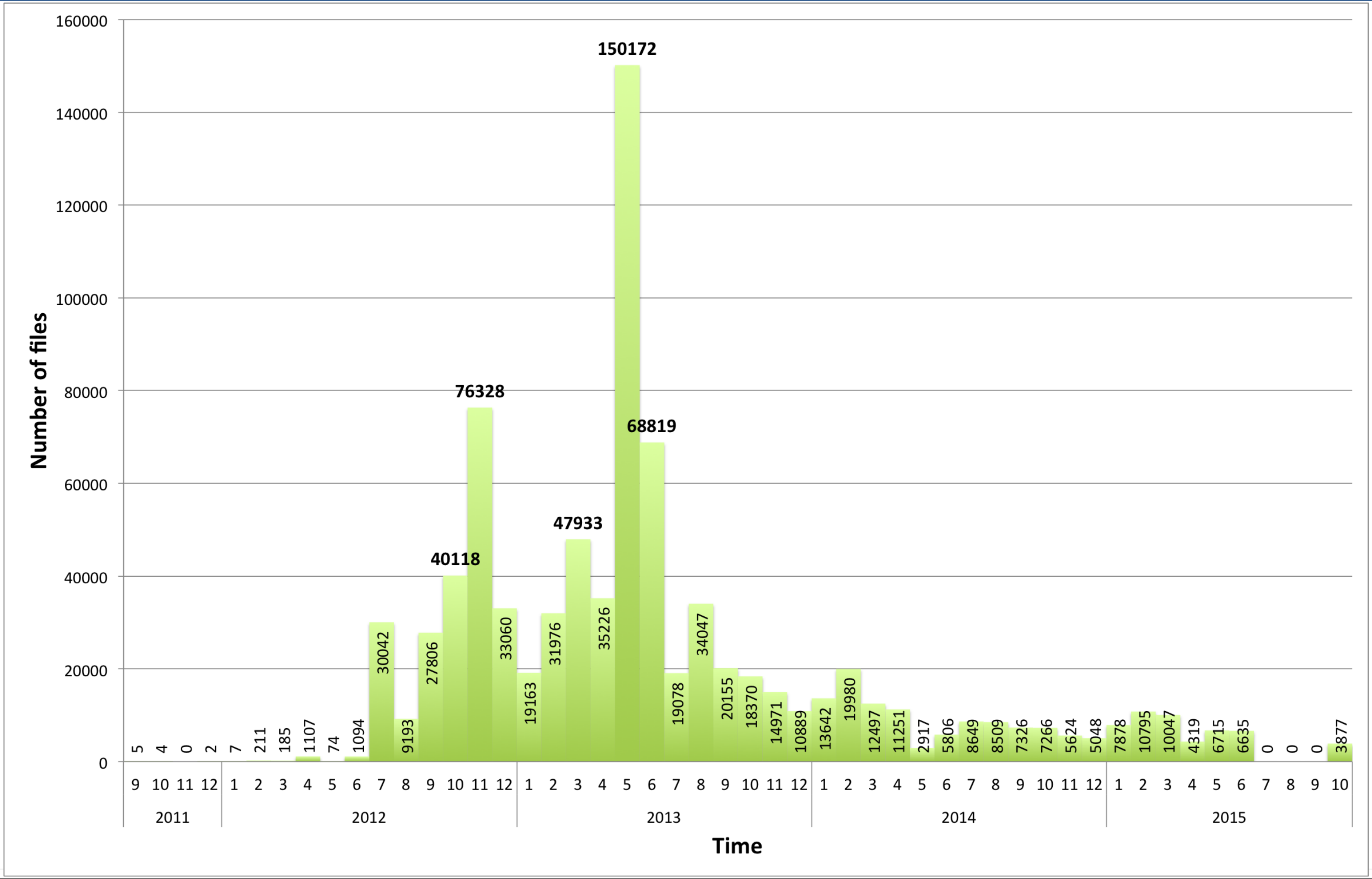
Downloaded data (GB) per month
(LLNL data node view)



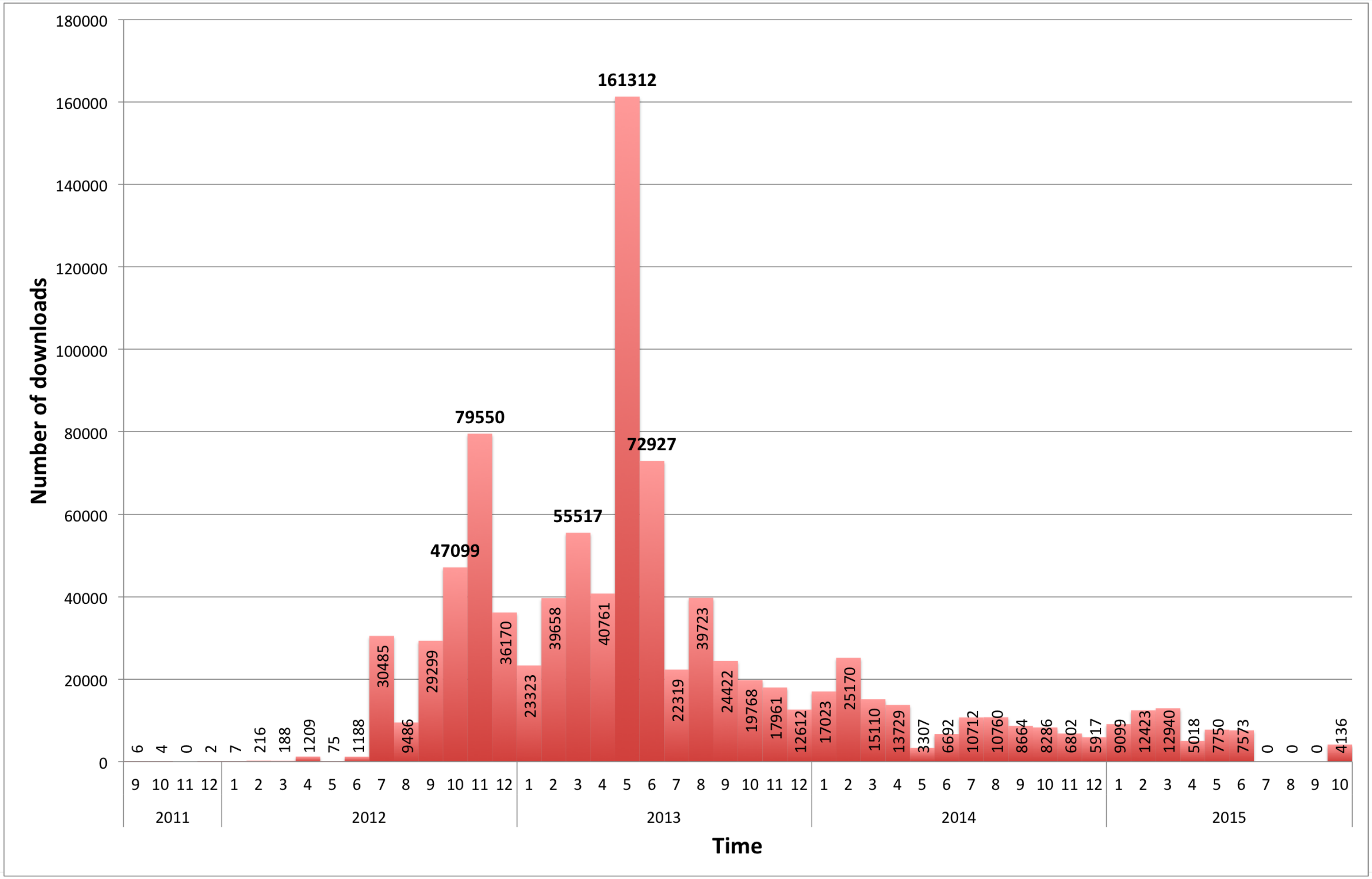
Number of distinct users per month
(LLNL data node view)



Number of distinct files per month
(LLNL data node view)

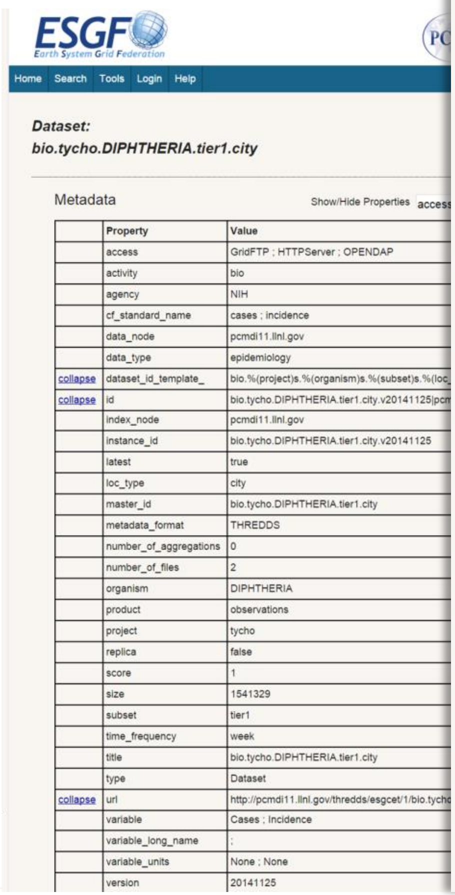


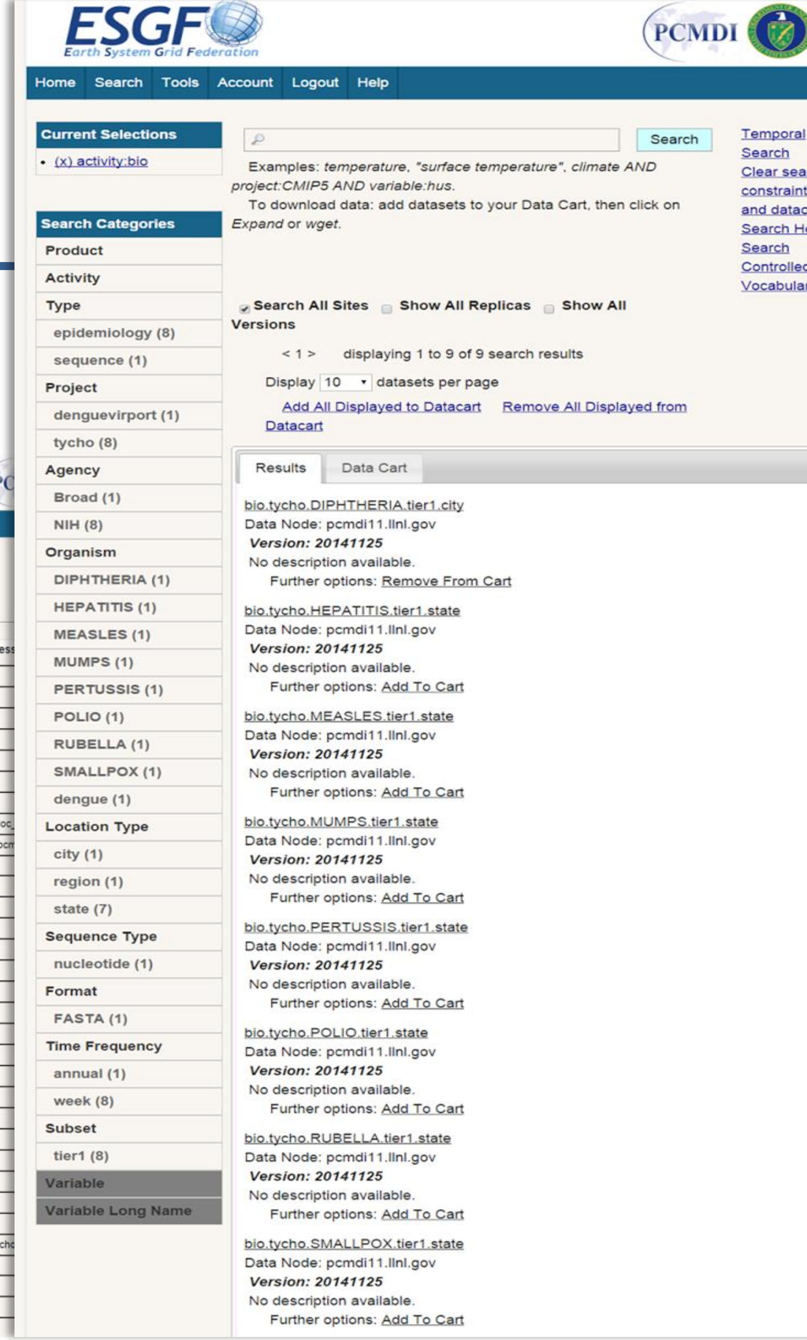
Number of downloads per month
(LLNL data node view)



ESGF can do much more than climate research

- What else can we use this large federated data platform for?
- Many other research communities need a platform that can distribute peta- and exascale data
  - Physics
  - Astronomy
  - Energy
  - Biology





Lawrence Livermore National Laboratory