LLNL-PRES-575652

Prepared by LLNL under Contract DE-AC52-07NA27344.

U.S. DEPARTMENT OF ENERGY | Office of Science

Earth System Grid Federation (ESGF): Future and Governance

October 2, 2012

Dean N. Williams

On behalf of Multiple Earth System Communities and Projects

ESGF Earth System Grid Federation

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

World Climate Research Programme (WCRP), Working Group on Coupled Modelling (WGCM)—Stakeholders and ESGF

Encourage the formalization of ESGF with international inter-agency agreements, orchestrated by WCRP.

- **WCRP and National Research Council (NRC) endorsement should be capitalized on by agencies to increase base funding for a global data infrastructure.**
- **Governance proposal being developed**
- **Infrastructure cannot be financed with soft money!**

All MIPS should follow the lead and standards set by CMIP5.

- **With that understanding, CMIP could be divided into smaller, more focused and manageable sets of experiments, which would be less disruptive to the scientific life at the centers.**

The ESGF software stack should be used to handle the data needs for other coordinated modeling activities (obs4MIPs, ana4MIPs, etc.).

A brief history of the Earth System Grid (ESG): ESG-I, ESG-II, ESG-CET, ESGF

ESG-I funded under DOE's Next Generation Internet (NGI) to address the emerging challenge of climate data 1999 – 2001 (ANL, LANL, LBNL, LLNL, NCAR, USC/ISI)

Data movement and replication; Prototype climate "data browser"; Hottest Infrastructure" Award at SC'2000.

ESG-II funded under DOE's Scientific Discovery through Advanced Computing (SciDAC), turning climate data sets into community resources 2001-2006 (ORNL addition)

Web-based portal, metadata, access to archival storage, security, operational services, 2004 first operational portal CCSM (NCAR), IPCC CMIP3/AR4 (LLNL); 200 TB of data, 4,000 users, 130 TB served.

ESG-CET funded under DOE's Offices of ASCR and BER to provide climate researchers worldwide with access to: data, information, models, analysis tools, and computational resources required to make sense of enormous climate simulation and observational data sets 2006 – 2011 (PMEL addition)

2010 Awarded by American Meteorological Society (AMS) for leadership which led to a new era in climate system analysis and understanding.

CMIP3, CMIP5, CCSM, POP, NARCCAP, C-LAMP, AIRS, MLS, Cloudsat, etc.

25,000 users, 500-800 users active per month, over 1 PB served

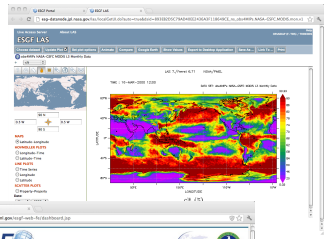
ESGF P2P under the DOE's Office of BER, it is an open consortium of institutions, laboratories and centers around the world that are dedicated to supporting research of climate change, and its environmental and societal impact. (Additional U.S. funding from NASA, NOAA, NSF.) The federation includes: multiple universities and institution partners in the U.S., Europe, Asia, and Australia.

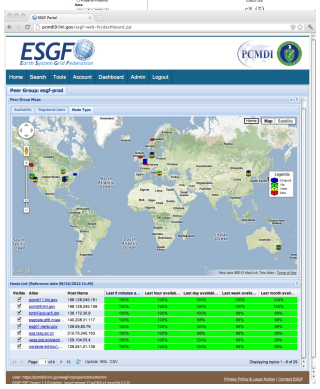
ESGF Data Holdings (1.8 PB)

Phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5)
Coordiated Regional climate Downscaling Experiment (CORDEX)
Climate Science for a Sustainable Energy Future (CSSEF)
European Union Cloud Intercomparison, Process Study & Evaluation Project (EUCLIPSE)
Geo-engineering Model Intercomparison Project (GeoMIP)
Land-Use and Climate, Identification of robust impacts (LUCID)
Paleoclimate Modeling Intercomparison Project (PMIP)
Transpose-Atmospheric Model Intercomparison Project (TAMIP)
Clouds and Cryosphere (cloud-cryo)
Observational products more accessible for coupled model intercomparison (obs4MIPs)
Reanalysis for the coupled model intercomparison (ANA4MIPs)
Dynamical Core Model Intercomparison Project (DCMIP)
Community Climate System Model (CCSM)
Parallel Ocean Program (POP)
North American Regional Climate Change Assessment Program (NARCCAP)
Carbon Land Model Intercomparison Project (C-LAMP)
Atmospheric Infrared Sounder (AIS)
Microwave Limb Sounder (MLS)

ESGF is more than CMIP: federated and integrated data from multiple sources

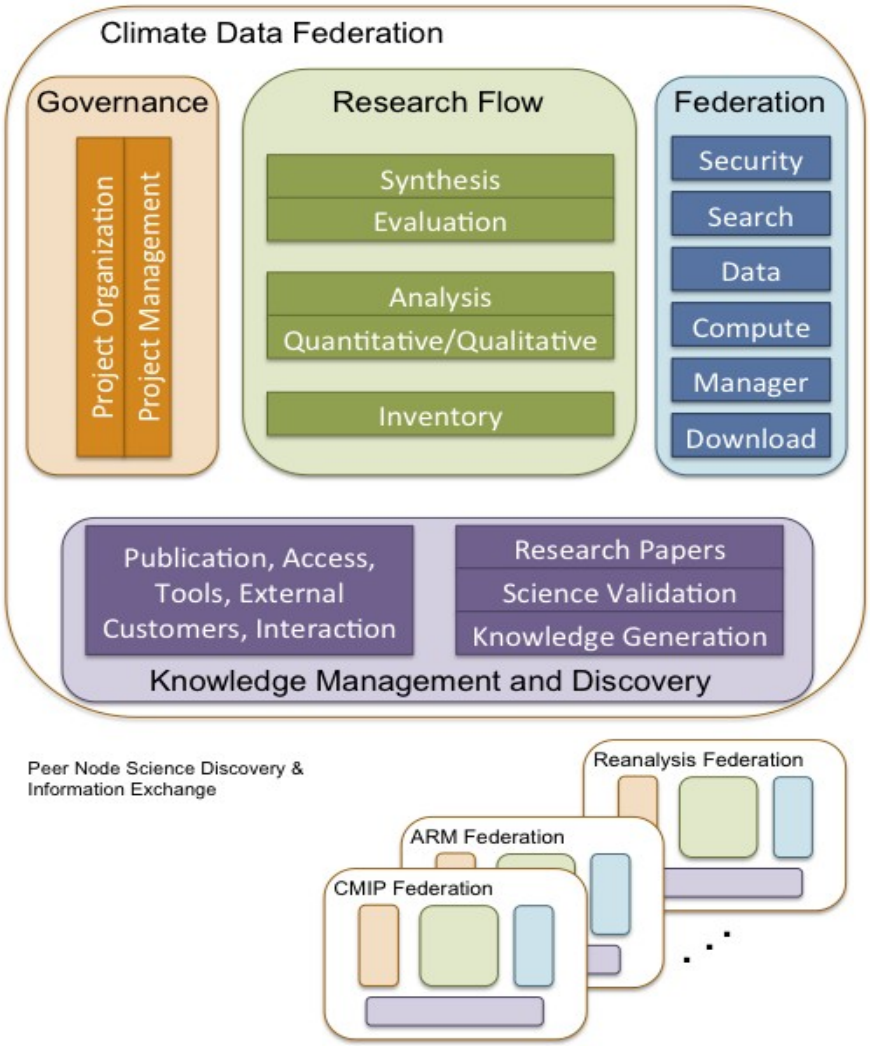
http://pcmdi9.llnl.gov/esgf-web-fe/





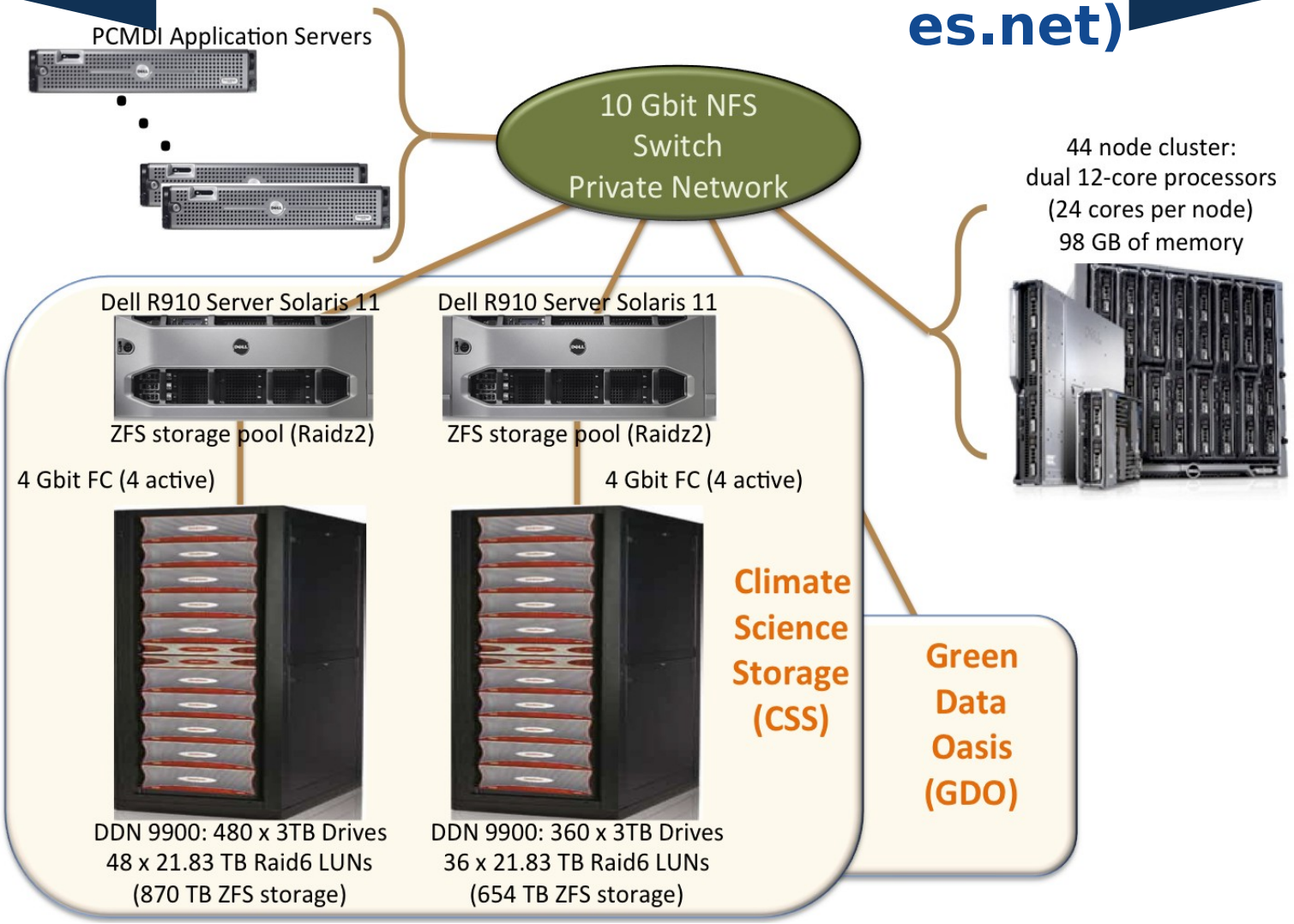
ESGF software system integrates data federation services

- NetCDF Climate and Forecast (CF) Metadata Convention
  - (LibCF)
  - Mosaic
- Climate Model Output Rewriter 2 (CMOR-2)
- Regridders: GRIDSPEC, SCRIP, & ESMF
- Publishing
- Search & Discovery
- Replication and Transport
  - GridFTP, OPeNDAP, DML, Globus Online, ftp, BeSTMan (HPSS) Networks
- Data Reference Syntax (DRS)
- Common Information Model (CIM)
- Quality Control
  - QC Level 1, QC Level 2, QC Level 3, Digital Object Identifiers (DOIs)
- Websites and Web Portal Development
  - Data, Metadata, Journal Publication Application
- Notifications, Monitoring, Metrics
- Security
- Product Services
  - Live Access Server, UV-CDAT



ASCR Climate 100 project: connecting our system to the largest network in the world and organizing the climate network community

(climatedata@es.net)



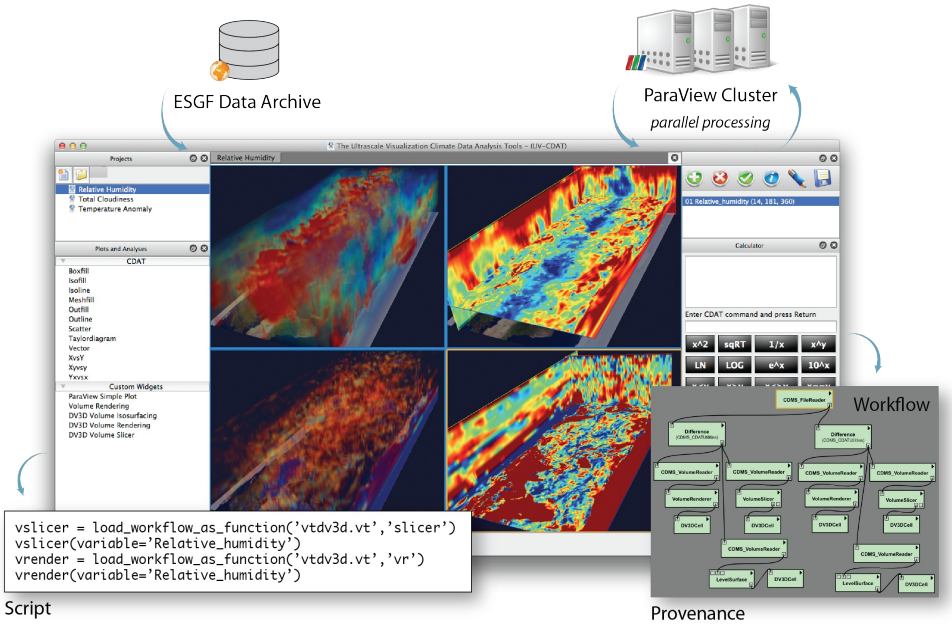
Supported data activities (complementary DOE funded project)

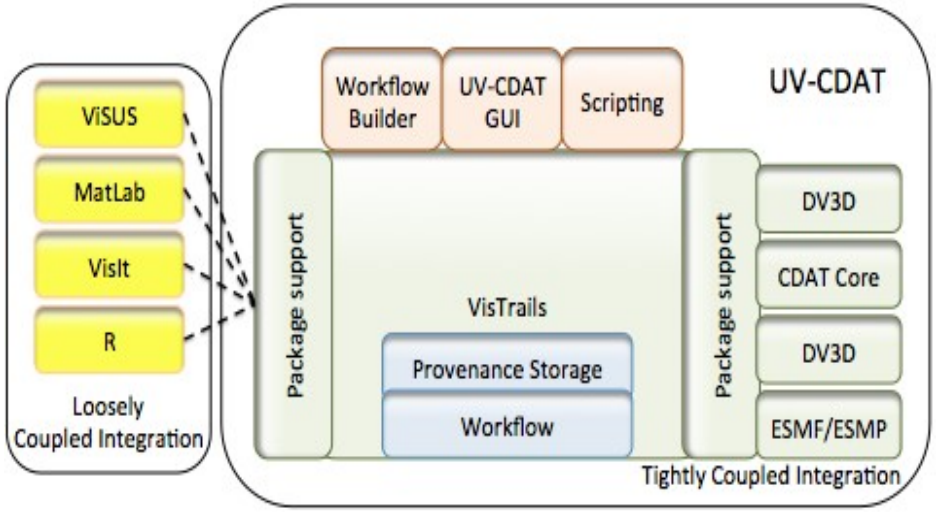
U.S. DEPARTMENT OF ENERGY | Office of Science

http://uvcdat.llnl.gov

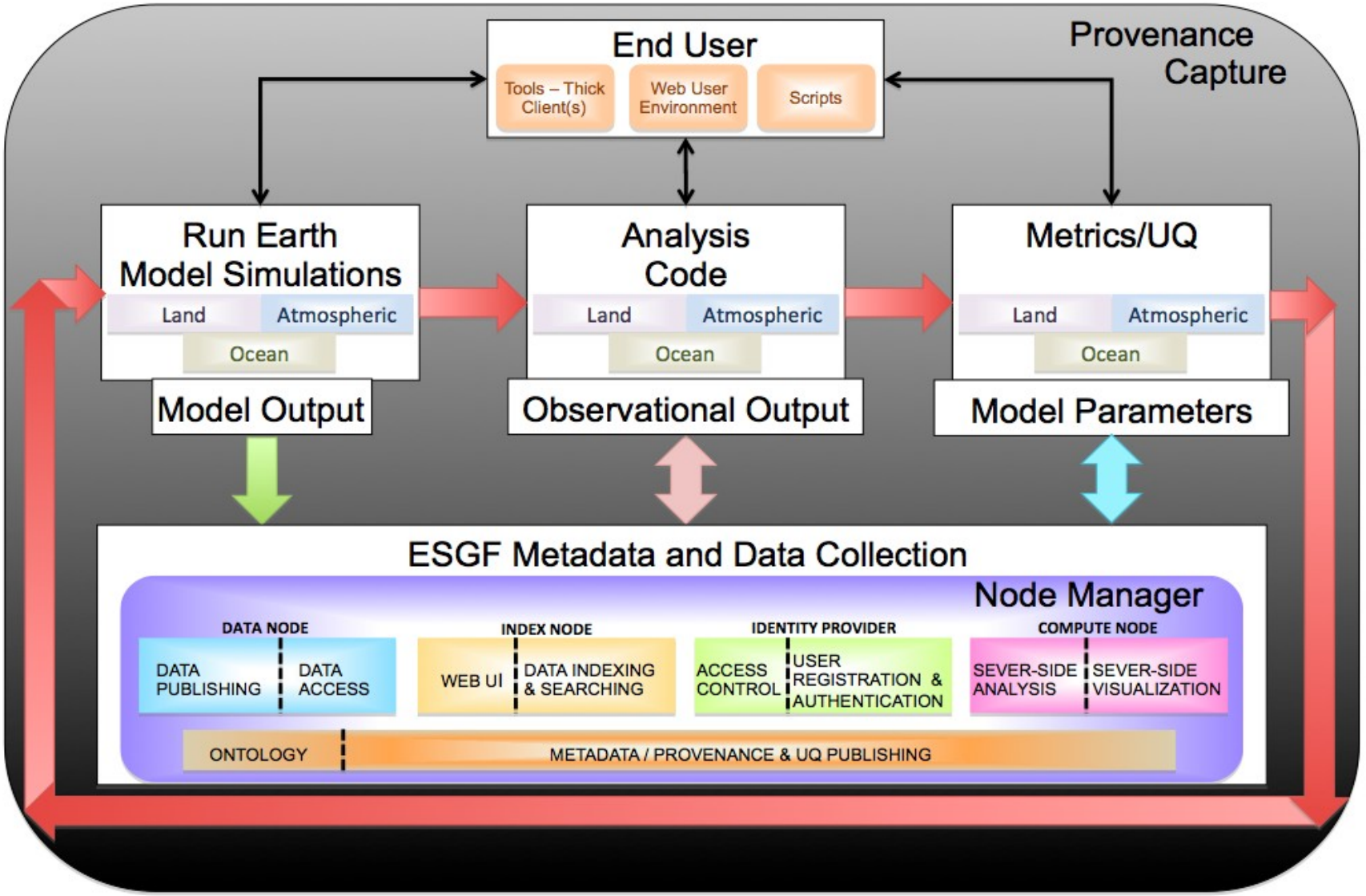
Ultra-scale Visualization Climate Data Analysis Tools (UV-CDAT):

- **Integrate** DOE's climate modeling and measurements archives
- Develop **infrastructure** for national and international model/data comparisons
- **Deploy** a wide-range of climate data visualization, diagnostic, and analysis tools with familiar interfaces for very large, high resolution climate data sets (**CDAT, VTK, R, VisIt, ParaView, DV3D**, …)
- **Workflow** – data flows are directed graphs describing computational tasks
- Takes advantage of **ESGF data management**

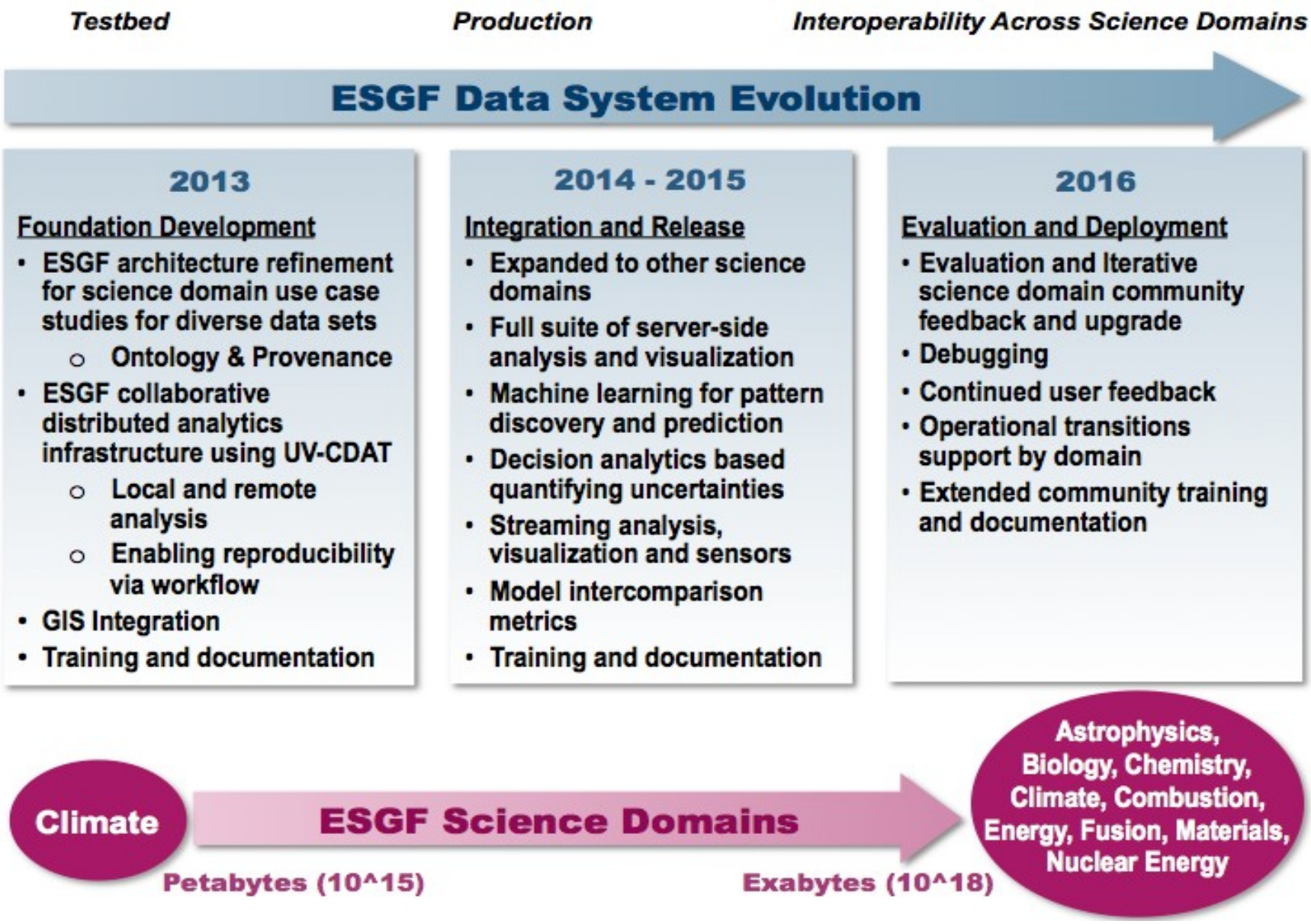




High-level Conceptual View of CSSEF Test Bed Architecture and Workflow



ESGF Roadmap: Vision for the Future



ESGF Governance

Governance – critical ideas of fairness, transparency, responsibility, accountability

The ESGF Review Board (ERB) consist of two committees:

Executive Committee

- **General guidance and overall high-level decisions in directing the course of the ESGF project in correspondence with multiple sponsor and stakeholders needs—ultimate responsibility for ensuring that ESGF meets the needs of customers and stakeholders.**

Technical Committee

- **General guidance and overall high-level decisions in directing the course of the ESGF project in correspondence with multiple sponsor and stakeholders needs—ultimate responsibility for ensuring that ESGF meets the needs of customers and stakeholders.**

ERB Committee Roles

Chair and/or co-Chairs

Organizes meeting agendas and maintains the roadmap and the list of outstanding proposals that require ERB intervention

Chair or co-Chairs may invite individuals or groups who have submitted proposals to present their plans at ERB meetings

Chair or co-Chairs are responsible for facilitating discussions and consensus building in their perspective committees

Voting Members

For each committee, there can be only one voting member per institution

Non-voting Members

Non-voting members on the executive committee must be voting members on the technical committee and non-voting members on the technical committee must be voting members on the executive committee.

Participate Members

Each committee can invite participants to join meetings. Participants are guests and hold no voting privileges. They may contribute to the code base and participate in meetings on limited bases.

ERB Intervention

Code changes with a high impact on developers and/or users are to be reviewed by the ERB committees.

Guiding principles for deciding whether a change require ERB involvement include:

- **Will the change significantly affect backwards compatibility?**
- **Will the change significantly affect users, managers, or developers?**
- **Does the change significantly shift the functionality and scope of ESGF?**
- **Are there any licensing issues?**

Executive Committee Responsibilities

ERB executive members are responsible for nominating new members, who are elected by consensus or majority vote (with the Chair breaking any tie)

Executive committee is responsible for organizing and interacting with the ESGF steering committee.

- **The steering committee—composing of sponsors, stakeholders, and users—provides input to the ERB executive committee on the entire ESGF project and path forward from their perspective**

Role of the executive committee:

- **Setting the strategic direction of the ESGF project**
- **Oversight of technical committee (and other committees as defined by the executive committee)**
- **Review of requirements (e.g., analysis, priorities, architecture, development, etc.) reported by the technical committee to the executive committee on a defined basis**
- **Approval of major architectural changes**

Technical Committee Responsibilities

ERB technical members are responsible for nominating new members (with final approval of the ERB executive committee) who are elected by consensus or majority vote (with the Co-Chairs breaking any tie)

Role of the technical committee:

- **Requirements analysis (interfaces with customers and users to understand requirements)**
- **Technical architecture**
- **Organizing development**
- **Testing and releases**
- **Reporting to the executive committee**

Meetings

ERB committees will hold a joint meeting on a schedule of their choosing and convenience but at least four times a year. The executive committee will meet monthly. The technical committee will meet weekly. The Chairs of both committees will communicate weekly to monthly.

Every 18 months, a face-to-face meeting shall occur to detail progress and future community requirements

Questions and discussion

Click to edit Master text styles

- Second level
  - Third level
    - Fourth level
      - Fifth level