

World Climate Research Program

WCRP Data Advisory Council

Updates on NOMADS, GEO and ESGF



April 7-8, 2016
The Collider
Asheville, NC

Glenn K. Rutledge

NOAA NCEI Asheville

NOMADS Principal Investigator

White House USGCRP Interagency Group on Modeling

GEO Climate Data Access Task Lead

Global Organization for Earth System Science Portals (GO-ESSP)

Founding member and former co-Chair NOAA's DMIT



Outline

- NOAA Operational Model Archive and Distribution System (NOMADS)
- Group on Earth Observations (GEO)
- Earth System Grid Federation (ESGF)
- Q&A



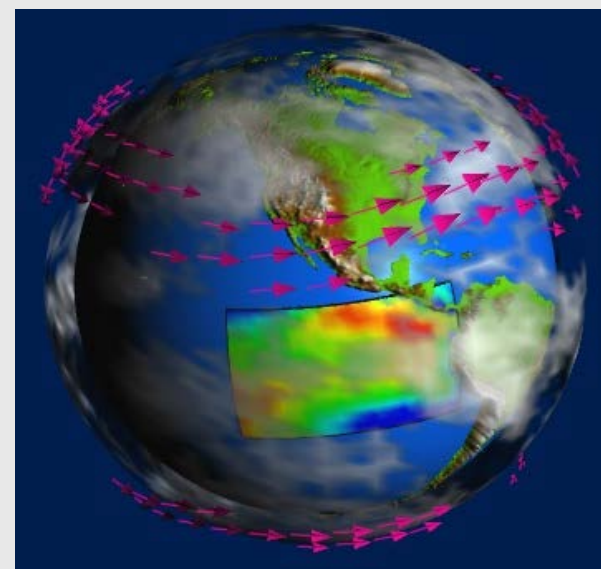
NOMADS

A Modeling Community Resource since 2002



- Until 2002 there existed no long-term archive for Climate and Weather models in NOAA.
- University and Institutional research went largely untapped by NOAA scientists. Effort was wasted on data receipt and format issues with no infrastructure to collaborate.

- Retrospective analysis and model inter-comparison are necessary to verify and improve short term NWP models, seasonal forecasts, climate assessments and detection





NOMADS Data Access System

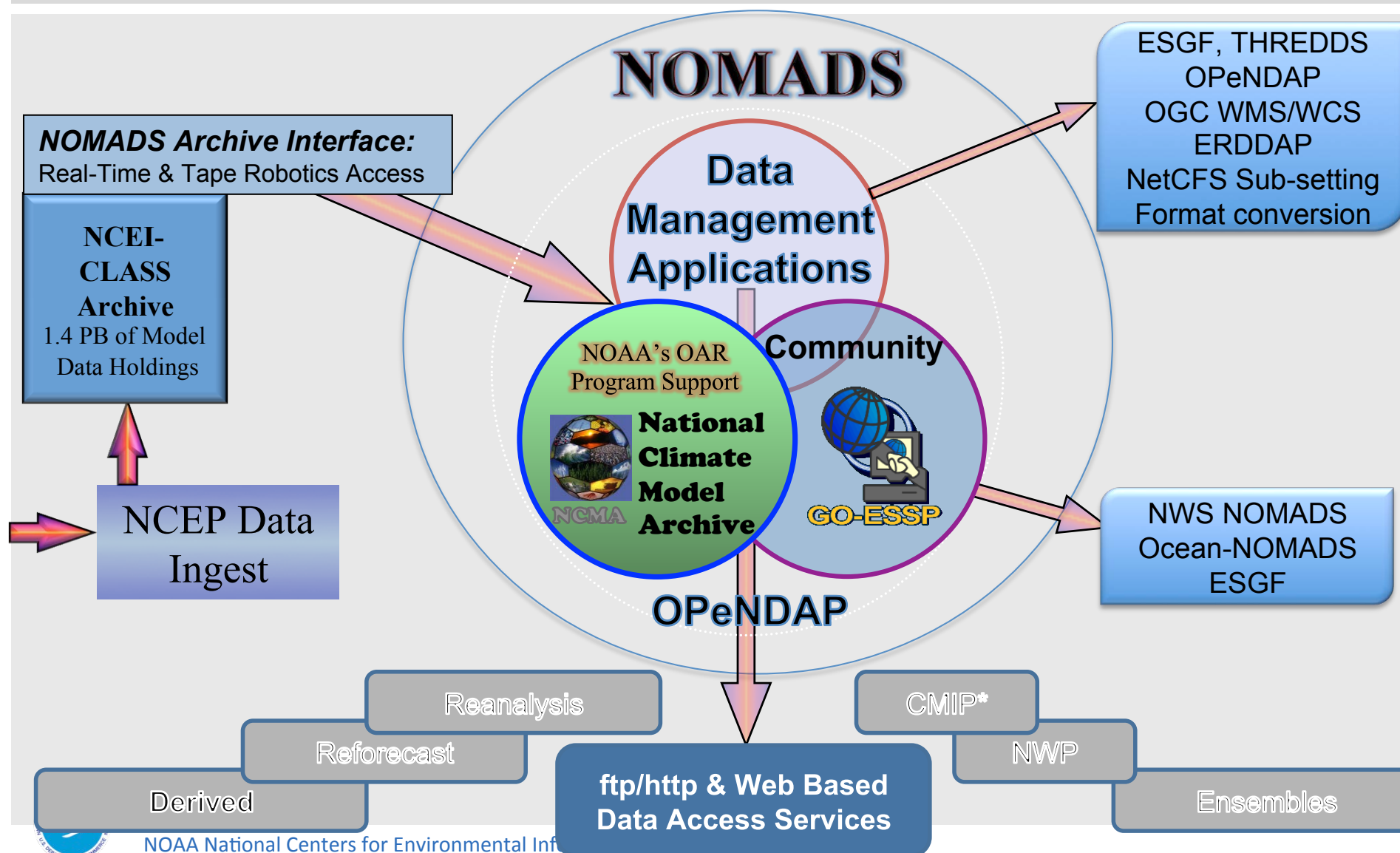
- To overcome this deficiency in model data access, some of the Nations top scientists actively engaged in a grass-roots framework to share data and research findings over the Internet.
- In 2001 NCDC, NCEP, and GFDL initiated the NOAA Operational Model Archive and Distribution System with PMEL, NCAR, COLA, PCMDI, ...
- Operational in NESDIS and NWS, NOMADS is a distributed data service providing access to climate and weather models. Forerunner to GO-ESSP



NOMADS

The NOAA Operational Model
Archive and Distribution System

NCEI NOMADS System



Ensemble Probability Tool

[Ensemble Probabilities](#) on the fly:

21 members

16 days of forecast

6 hrly increments

Restful based PDF's on the fly!

The NOMADS Ensemble Probability Tool is a tool that is designed to allow users to interrogate the NCEP Global Ensemble model. The tool allows the user to describe a set of conditions and determine the probability that that set of conditions will occur at a given location.

The NOMADS Ensemble Probability Tool queries the 21 member GFS ensemble dataset located on the NCEP NOMADS High Availability server. The data is passed via OpenDAP back to the application, where it is read using the Java NetCDF library, and then the probabilities are calculated.

For more information, please see our [help page](#).

Where

☒ Station ID
☐ Lat Lon

When

☒ Latest model run (2009 Oct. 23 06z)
☐ Year Month Date Model Run

What

☒ Air Temperature at 2 meter height
☒ 6 hour Highest temperature
☐ 6 hour Lowest temperature
☐ Precipitation
☐ Wind at 10 meter height
☐ Cloud Cover
☐ Air Temperature at 850 millibar pressure level
☐ Convective Available Potential Energy (CAPE)

Submit

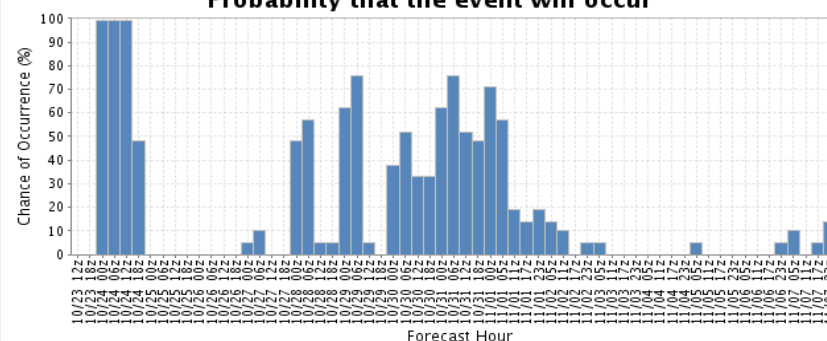
Request

Location:
Asheville, Asheville Regional Airport, NC, United States (35-25-55N, 082-32-15W)
Time:
Oct 23, 2009 06z
Event:
where the highest temperature is greater than 65 degrees F.

Progress [show]

Results

Probability that the event will occur



NOMADS Ensemble Probability Tool

Probability that the following event will occur:
where the highest temperature is greater than 65 degrees F.

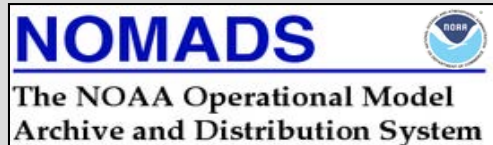
At the location:
Asheville, Asheville Regional Airport, NC, United States (35-25-55N, 082-32

For the GENS model run at the given time:
Oct 23, 2009 06z

<http://nomads.ncdc.noaa.gov/EnsProb/>

NOAA National Centers for Environmental Information





NCEI NOMADS Data Holdings

NWP Model and Climate Reanalysis:

1.4 PB of model data holdings as of 2016

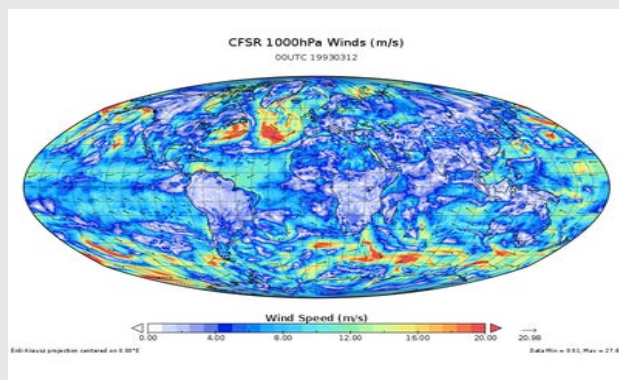
- Global Forecast System (GFS), 1 and ½ degree Mar 2004 – Present
- NCEP Climate Forecast System Reanalysis (CFSRR) Global 0.5 deg Jan 1979 – Mar 2011
- NCEP Climate Forecast System Reforecast (CFSRR) Global 0.5deg Jan 1979 – Mar 2011
- NCEP North American Regional Reanalysis (NARR) 30 years 32km Jan 1981 – Present
- NCEP/NCAR/DOE NNR2 Global Reanalysis Jan 1979 – Jan 2005
- NCEP Global Ensembles (GENS) / TIGGE (w/ ECMWF) Dec 2007– Present
- **NCEP Spectral Statistical Interpolation (SSI) Global Data**
- **Assimilation System (GDAS) w/ model restart Feb 2002 – Present**
- North American Mesoscale (NAM, formally Eta) 12km Feb 2005 – Present
- Rapid Update Cycle (RUC) / RAP 20km and 13km Jan 2006 – Present

2015 usage:

- Successful requests: 248 million
- Distinct hosts served: 136,193
- Data transferred: 1.29 petabytes



NOAA's Climate Forecast System Reanalysis and Reforecast



A global plot of CFSR near-surface winds on March 12, 1993, at 00 UTC.

- In 2010 the Environmental Modeling Center at NOAA/ NWS/NCEP, produced a 30-yr Reanalysis that included 6-hrly Reforecasts.
- The CFS is a fully coupled model representing the global interaction between Earth's oceans, land, and atmosphere.
- Hourly global data down to one-half degree resolution.
- Access to the 879TB dataset is provided solely by NCEI/NOMADS over the following time periods:

Reanalysis - January 1979 to March 2011

Reforecasts - December 1981 to March 2011

CFSv2 Operational Analysis and Forecasts - April 2011 to Present

URL

<http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/climate-forecast-system-version2-cfsv2>

GEO Vision

*To realize a future wherein decisions and actions,
for the benefit of humankind, are informed by
coordinated, comprehensive and sustained
Earth observations and information*



prepared by André Obregón,
GEO Technical Expert for Climate



GEO-XII Plenary & Ministerial Summit Mexico City 9-13 November 2015

- Adoption of a ten year Strategic Plan (2016 - 2025)
- Ministerial Declaration that focuses on harnessing critical environmental observations to enable leaders to make better-informed decisions for the benefit of humanity at a time of rapid global change

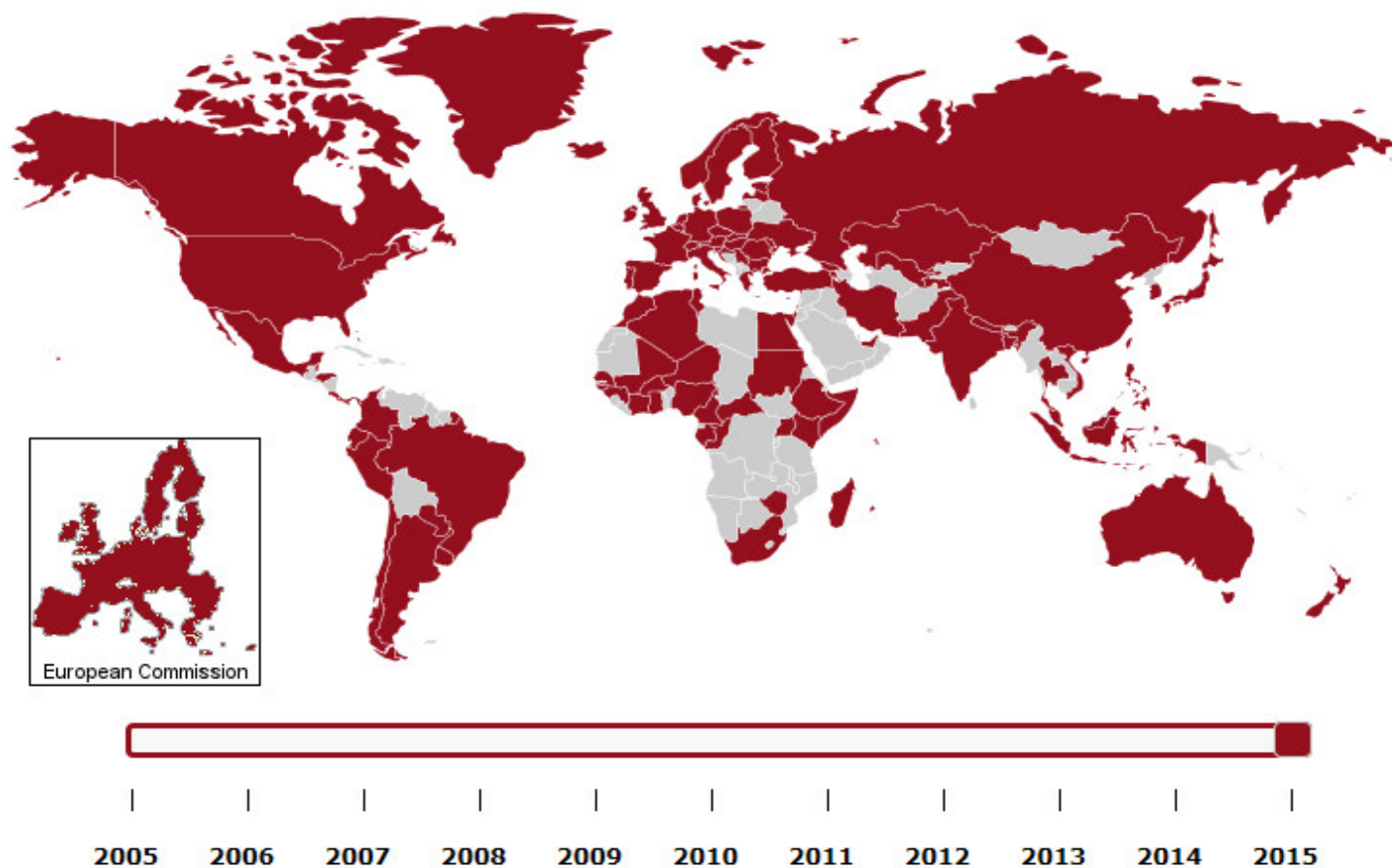
GEO Objectives

- **Improve and coordinate observation systems globally**
- **Provide access to data & information**
- **Advance broad open data policies/practices**
- **Build capacity**



102 GEO Members

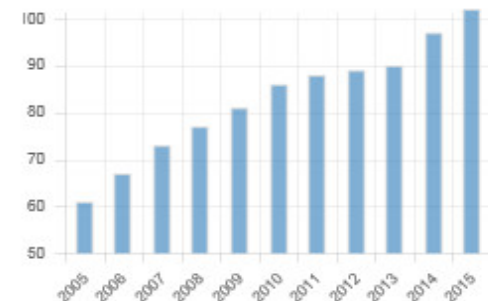
GEO Member Map for the year 2015



Number of Members (2015)

Africa:	27
Americas:	16
Asia/Oceania:	18
C.I.S.:	7
Europe:	34
Total:	102

Number of Members by year



© 2013 Pearson Education, Inc. or its affiliate(s). All rights reserved. This publication is protected by copyright. Any unauthorized distribution or reproduction of this work is illegal. All other rights reserved.



Unique Time in GEO's History

- Transition to next decade 2016-2025
- Recognition of GEO's convening power – Members, POs, Development Banks, Foundations, emerging Private Sector
- Evolution & Recognition of Policy Mandates
- New Strategic Plan with new programmatic mechanisms – community activities, foundational tasks, initiatives and flagships



Strategic Plan emphasizes climate change and its impacts as a cross-cutting area

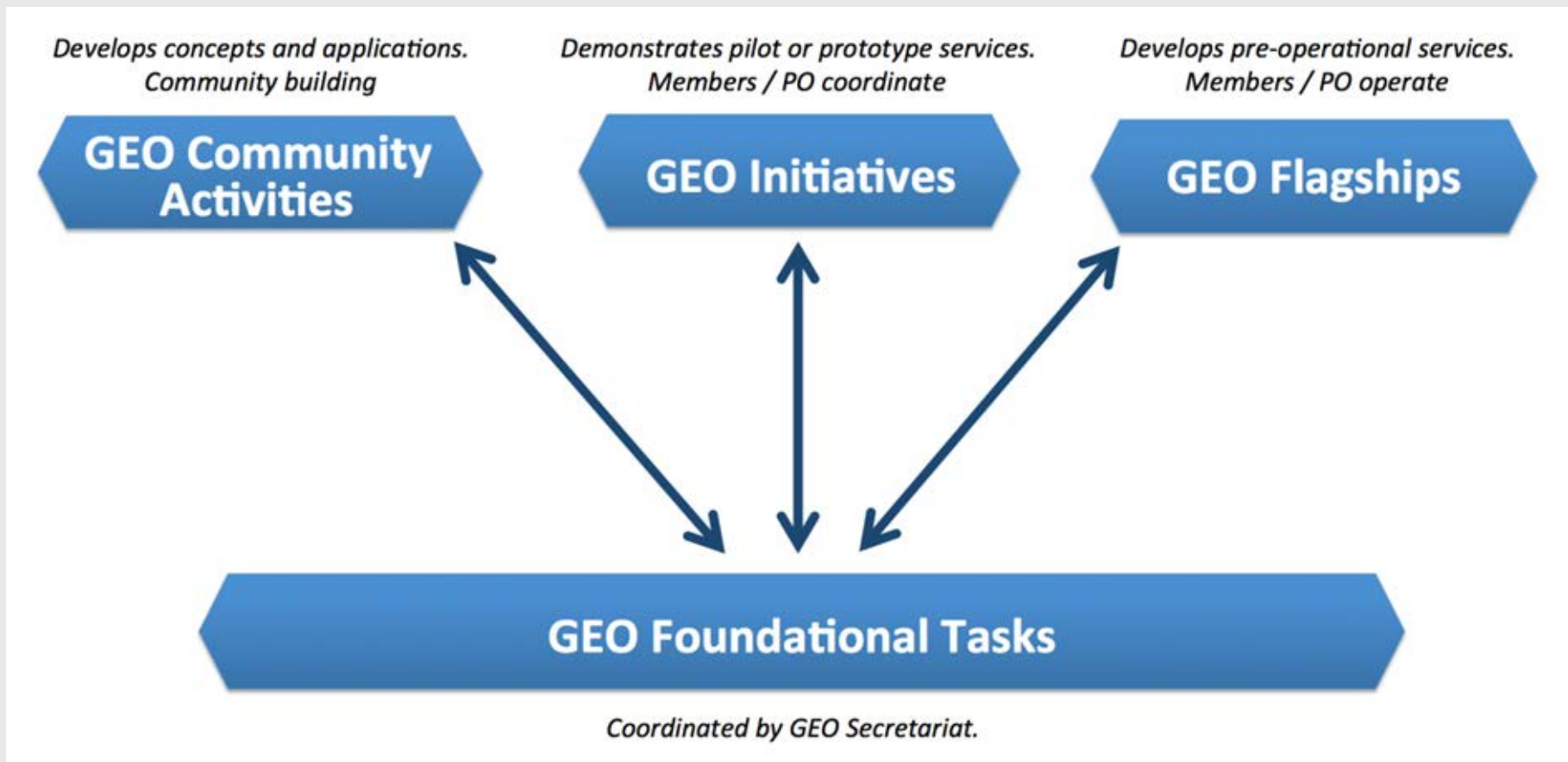
ipcc
climate change AR5 WG II

Sectoral impacts

Freshwater resources
Terrestrial and freshwater ecosystems
Coastal systems and low-lying areas
Marine systems
Food security and food production systems
Urban and rural areas
Economic sectors and services
Human health and security
Livelihoods and poverty



Implementation Mechanisms



GEO Community Activity:

CA-03 “Access to climate data in GEOSS”

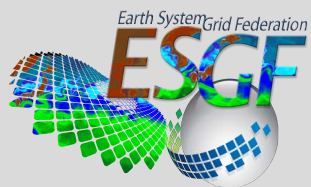
- **Some objectives:**

- Assist users of all levels of expertise to find and **discover** climate and weather information including **access** tools and **visualization**
- Foster development of **shared infrastructures** and on-line services within GEOSS and GCI.
- Leverage **CMIP** infrastructure for access to climate model outputs under the **ESGF** international collaboration
- Access to associated observational data sets under the **ESGF/obs4MIPs/ana4MIPs** activities

- **Contributors:**

- USA (NOAA, NASA), WRCP, CORDEX...





The Earth System Grid Federation (ESGF)

- ESGF is an easy to use and secure federated web based software data infrastructure for large climate datasets
- A virtual collaborative environment for diverse research and analysis tasks with large and varied datasets
- NCEI / CICS-NC ESGF “Pilot” in progress for a developmental node to advance research interests in support of CMIP-5 and CMIP-6



<http://esgf.llnl.gov>





ESGF¹ is a coordinated multiagency, international collaboration of institutions that continually develop, deploy, and maintain software needed to facilitate and empower the study of climate

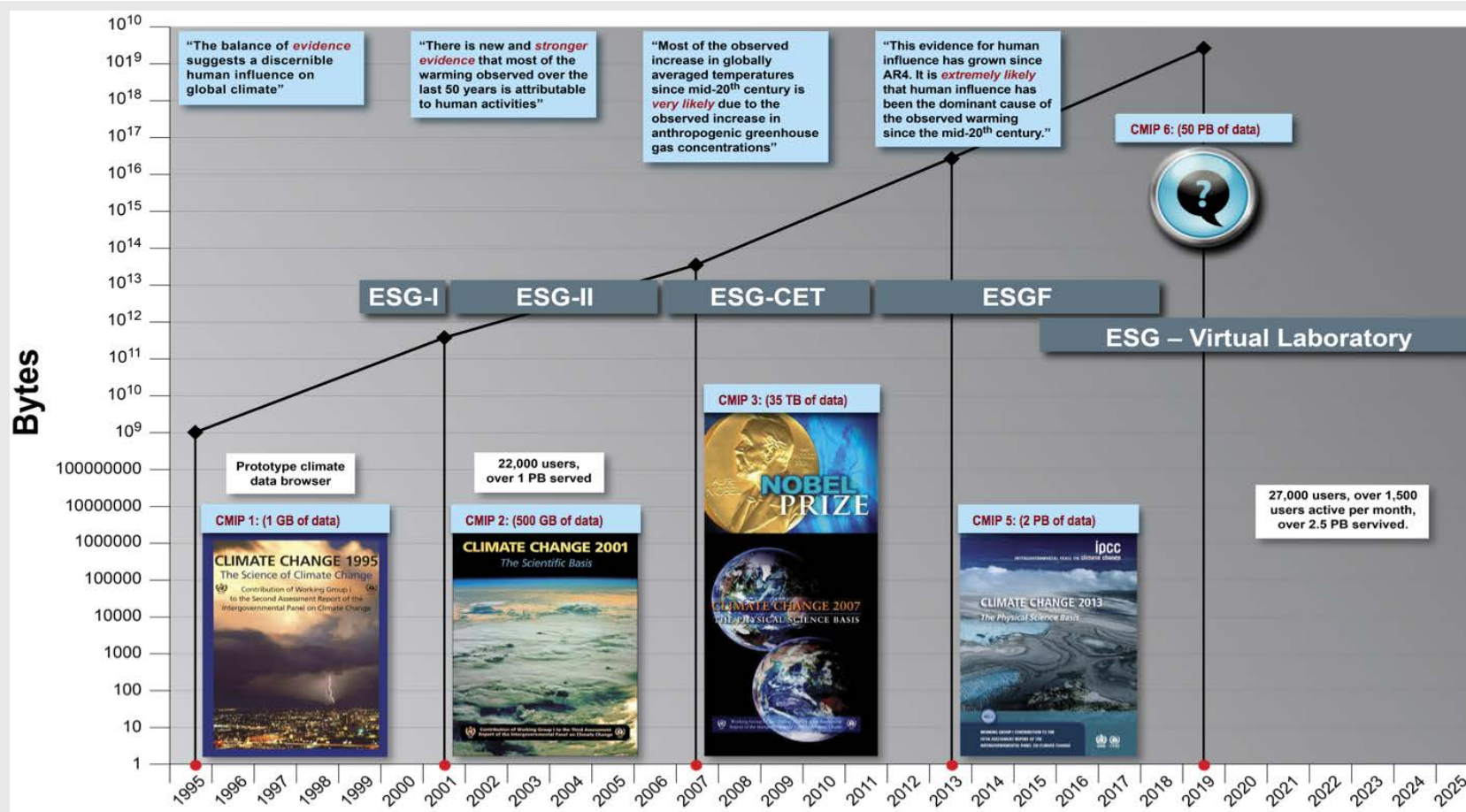


¹. Dean N. Williams, V. Balaji, Luca Cinquini, Sébastien Denvil, Daniel Duffy, Ben Evans, Robert Ferraro, Rose Hansen, Michael Lautenschlager, and Claire Trenham, "A Global Repository for Planet-Sized Experiments and Observations", Bulletin of the American Meteorological Society, early release, 2016,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





CMIP and ESGF history: scientific challenges and motivation use case



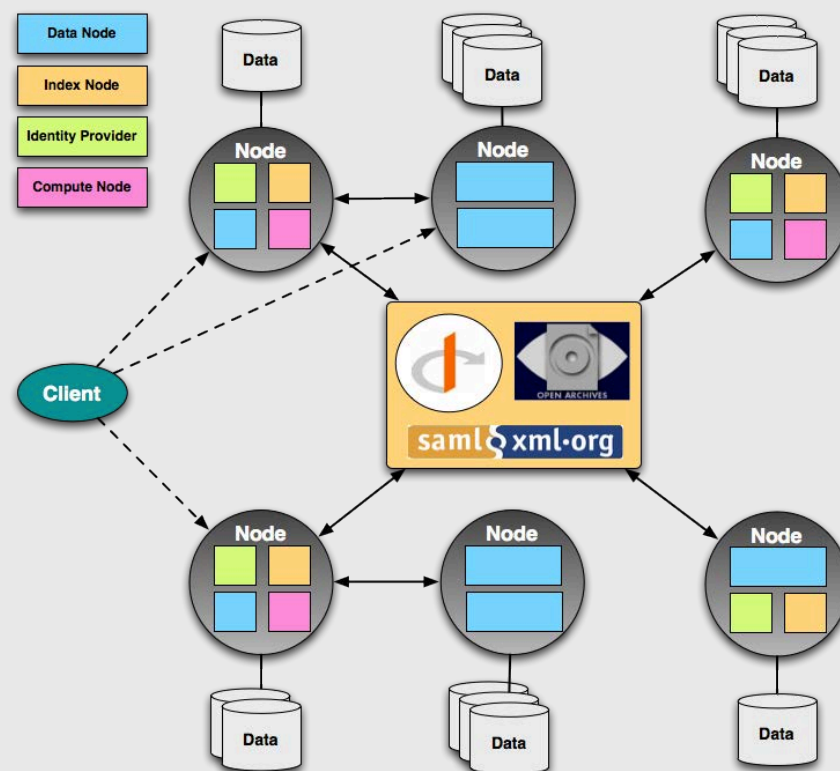
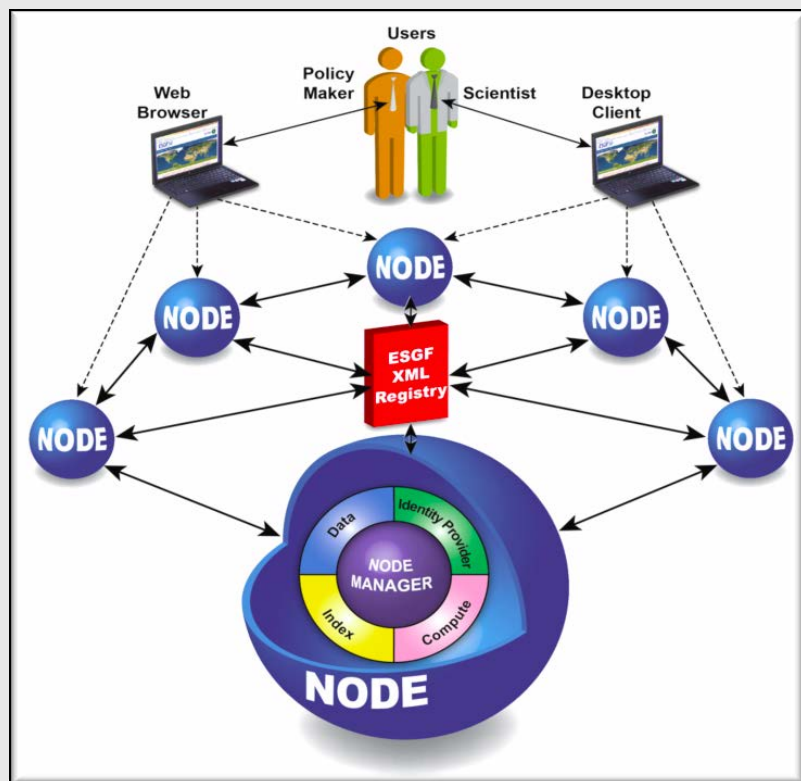
ESG recognizes that data management, stewardship and curation is an ongoing and long-lived function that requires a strategy that is resilient to continuing evolution in hardware and software.





ESGF software infrastructure

ESGF is a software infrastructure for management, dissemination, and analysis of simulation and observational data. The software utilizes hardware, networks, software for data management, access and processing.



ESGF federation nodes interact as equals. Users log onto any node using single sign-on OpenID to obtain and access data throughout the entire federation.

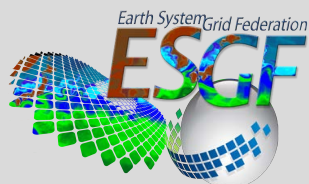


ESGF release version 2.0 (overhauled)

Following a security incident in June 2015, the ESGF system was brought offline and the software stack was extensively re-engineered to accomplish the following goals:

- Execute complete software scan of all modules, fix all exposed and other potential security breaches (NASA lead)
- Major upgrade of underlying system infrastructure:
 - CentOS7, Java 1.8, Tomcat 8, Postgres 9.4, OpenSSL 1.0, Python 2.7.9
 - Switch ESGF installer to RPM-based components
 - Run Apache httpd server in front of Tomcat (better performance, flexibility)
- Major upgrade of all ESGF services:
 - Search services (Solr5), data download (TDS5), high performance data transfer (Globus-Connect-Server), computation (UV-CDAT), visualization (LAS)
 - Replace old web-front-end with new CoG user interface
- Republish ALL data collections (CMIP5, CORDEX, Obs4MIPs, ana4MIPs, ...)





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>





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

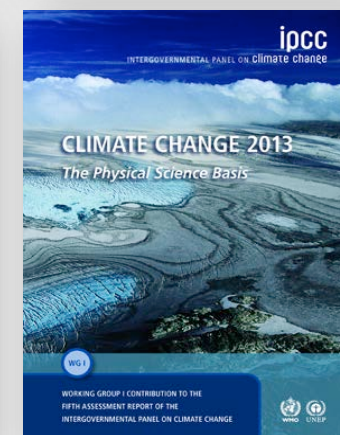
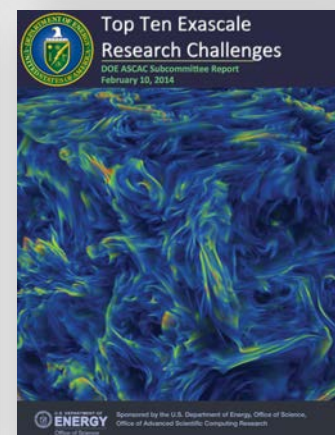
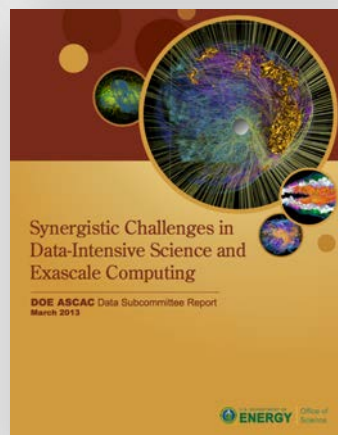
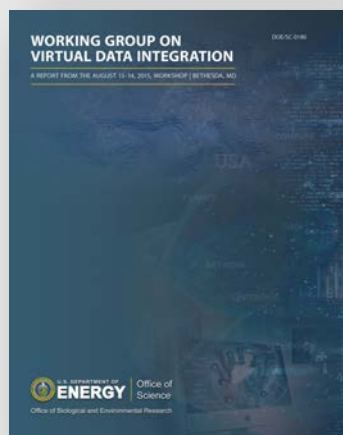
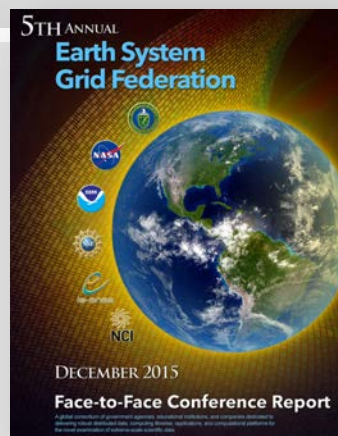
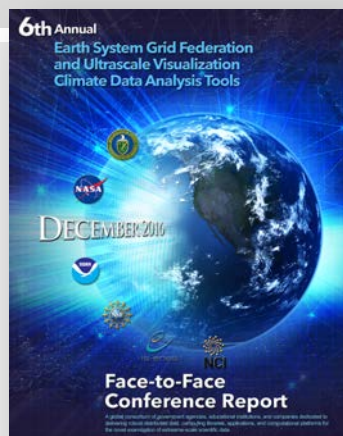
Immediate goal: 4 Gbps (1 PB/month) of sustained disk-to-disk data transfer between ESGF primary data centers

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





For More Information: Data Workshop and Conferences Reports



DOE ESGF workshop and conference reports can be found at:
<http://esgf.llnl.gov/reports.html>



- esgf.llnl.gov; ESGF public website
- esgf.llnl.gov/reports.html; ESGF reports
- uvcdat.llnl.gov; UV-CDAT public website
- icnwg.llnl.gov; international network website
- github.com/esgf; ESGF software repository website
- github.com/uv-cdat; UV-CDAT software repository website



NOAA/NCEI

NOMADS

GEO

ESGF

Questions?



Backup



Participants: NWS NOMADS

Real-Time NOMADS Access Services:

- NOMADS servers at the NWS provides real-time distribution including to NWS field offices.
- PB's transmitted yearly.

National Weather Service
NCEP Central Operations

Home News Organization

NOMADS
NOAA Operational Model Archive and Distribution System

Description of NOAA's NOMADS servers hosting NCEP model data

Users of NOMADS are reminded that they should use the URL <http://nomads.ncep.noaa.gov> to access the system and they will always be placed on the current active server. As of Tuesday October 7, 2009, users that have been using direct IP addresses to access NOMADS systems may no longer be able to access the system.

Help Desk: Questions or problems please use the link to submit a service ticket.
Background: Background documents about the NOMADS project.
Service Description: OCWWS Service Description Document

Click on link in the Data Set field for description and availability info.
Click on the column headings for description of each data access method.

Data Set	freq	grib filter	http	gds-alt
Global Models				
FNL	6 hours	grib filter	http	OpenDAP-alt
GFS 1.0x1.0 Degree	6 hours	grib filter	http	OpenDAP-alt
GFS 0.5x0.5 Degree	6 hours	grib filter	http	OpenDAP-alt
GFS 2.5x2.5 Degree	12 hours	grib filter	http	OpenDAP-alt
GFS Ensemble high resolution	6 hours	grib filter	http	OpenDAP-alt
GFS Ensemble Precip Bias-Corrected	daily	grib filter	http	OpenDAP-alt
GFS Ensemble high-resolution Bias-Corrected	6 hours	grib filter	http	OpenDAP-alt
GFS Ensemble NDGD resolution Bias-Corrected	6 hours	grib filter	http	OpenDAP-alt
NAEFS high resolution Bias-Corrected	6 hours	grib filter	http	OpenDAP-alt
NAEFS NDGD resolution Bias-Corrected	6 hours	grib filter	http	OpenDAP-alt
Regional Models				
AQM Daily Maximum	06Z, 12Z	grib filter	http	OpenDAP-alt
AQM Hourly Surface Ozone	06Z, 12Z	grib filter	http	OpenDAP-alt
HIRES Alaska	daily	grib filter	http	OpenDAP-alt
HIRES East CONUS	12 hours	grib filter	http	OpenDAP-alt
HIRES Guam	12 hours	grib filter	http	OpenDAP-alt
HIRES Hawaii	12 hours	grib filter	http	OpenDAP-alt
HIRES Puerto Rico	12 hours	grib filter	http	OpenDAP-alt
HIRES West CONUS	daily	grib filter	http	OpenDAP-alt
NAM-12 Alaska	6 hours	grib filter	http	OpenDAP-alt
NAM-12 CONUS	6 hours	grib filter	http	OpenDAP-alt

<http://nomads.ncep.noaa.gov/>

Participants: GFDL NOMADS

IPCC CMIP

List of data tables:

The table names listed below are consistent with those of the IPCC Standard

Table Amon:	Monthly Mean Atmospheric Fields and Some Surface Fields
Table aerg:	Monthly Mean Aerosol-Related Fields
Table cfDay:	CFMIP Daily-Mean Cloud Diagnostic Fields
Table cfMon:	CFMIP Monthly-Mean Cloud Diagnostic Fields
Table cfOff:	CFMIP monthly offline" Cloud Diagnostic Fields
Table cfSites:	CFMIP high frequency Cloud Diagnostic Fields
Table cf3hr:	CFMIP 3-Hourly Cloud Diagnostic Fields
Table day:	Daily Mean Atmosphere, Ocean and Surface Fields
Table fx:	Time-Invariant Fields
Table LImon:	Monthly Mean Land Cryosphere Fields
Table Lmon:	Monthly Mean Land Fields, Including Physical, Vegetation, Soil, a
Table Oclim:	Monthly Mean Ocean Climatology Fields
Table GOLD Oclim:	GOLD Monthly Mean Ocean Climatology Fields
Table OImon:	Monthly Mean Ocean Cryosphere Fields
Table Omon:	Monthly Mean Ocean Fields, Including Biogeochemical Fields
Table GOLD Omon:	GOLD Monthly Mean Ocean Fields
Table Ovr:	Annual Mean Ocean Fields, Including Biogeochemical Fields
Table 3hr:	2-D Atmospheric and Surface Fields Sampled Every 3 Hours
Table 6hrLev:	Fields (Sampled Every 6 Hours) for Driving Regional Models
Table 6hrPlay:	Fields (Sampled Every 6 Hours) for Storm-Track Analysis and
Problematic Variables: Issues with IPCC variables are addressed here.	



geophysical fluid
dynamics laboratory

[About us](#) [Research](#) [Products and Services](#) [Reference](#) [GFDL Only](#)

Spotlight on NOMADS

NOMADS is being developed as a Unified Climate and Weather Archive to provide Web access to model information so that users can make decisions about their specific needs. This spans time scales from days (weather), to months (El Nino), to decades (global warming). For more, see NCDC's [nomads](#) page.

Spotlight on GO-ESSP

The Global Organization for Earth System Science Portal (GO-ESSP) is a collaboration designed to build the infrastructure needed to create web portals to provide observed and simulated data to the climate and weather communities. The infrastructure created within GO-ESSP will provide a flexible framework of front-end and back-end software components. For more, see the [GO-ESSP](#) home page.

[gfdl's home page](#) > [products and services](#) > data portal

gfdl's data portal

CMIP5

<http://nomads.gfdl.noaa.gov/>

This section comprises of some information for CMIP5 project.

- [List of IPCC AR5 variables](#)
- [Available data](#)

GFDL CM2.0 and CM2.1 Climate Experiments:

Model output and documentation from a set of multi-century experiments performed using GFDL Intergovernmental Panel on Climate Change's 4th Assessment Report ([IPCC](#)) and the US Climate

For more information about the models and how to download their output, please refer to the [G](#)

Our Data Portal Services

Public data sets from GFDL are made available through the GFDL Data Portal. The Data Portal displays data file attributes that provide information about file contents without having to download

Our Public Data Files

Registration for the GFDL Data Files is free. Users are requested to complete the Registration Form when additional public data becomes available and when corrections are made to existing public

Data Storage

The data files on the data portal are stored in netCDF (network Common Data Form), and can be downloaded in a machine independent format for representing array-oriented scientific data, can be found at the

Software

This site implements the following:

- the Distributed Oceanographic Data System ([DODS](#)) package from [Unidata](#)/ [UCAR](#)
- the Grid Analysis and Display System - DODS Server ([GDS](#)) package from [COLA](#)/ [IGES](#)
- the Live Access Server ([LAS](#)) package from [PMEL](#)/ [NOAA](#).




NOMADS

The NOAA Operational Model
Archive and Distribution System



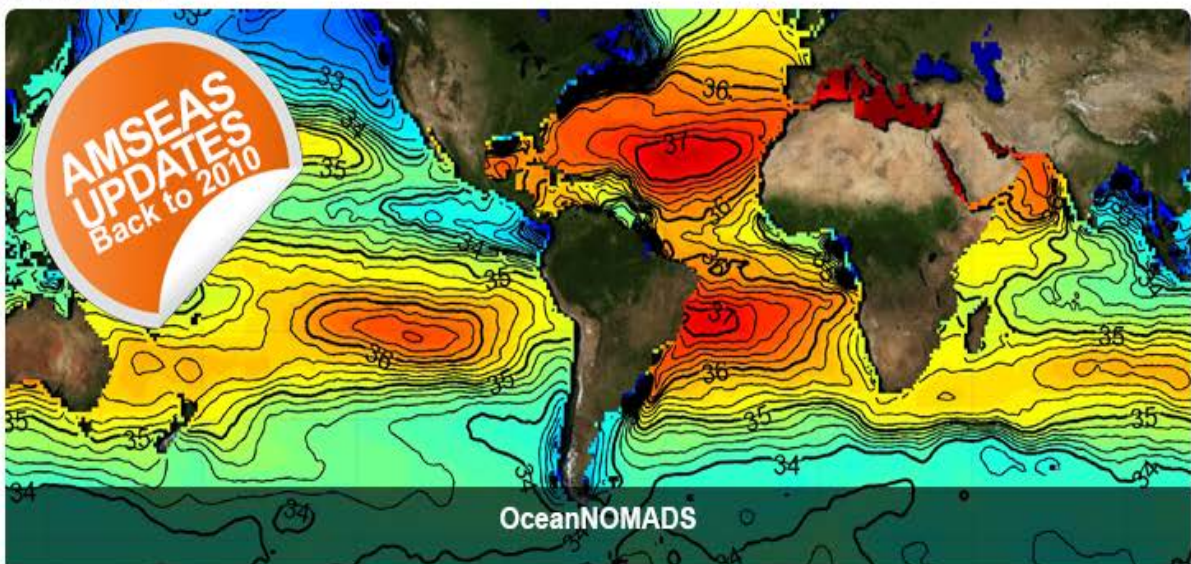
Participants: Ocean-NOMADS

[NOAA HOME](#) [WEATHER](#) [OCEANS](#) [FISHERIES](#) [CHARTING](#) [SATELLITES](#) [CLIMATE](#) [RESEARCH](#) [COASTS](#) [CAREERS](#)

 **NOAA** OceanNOMADS
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

[Home](#)
[NCOM](#)
[HYCOM](#)
[AMSEAS](#)
[USEAST](#)
[RTOFS](#)
[CFSR](#)

you are here: home



<https://www.ncdc.noaa.gov/data-access/model-data/ocean-nomads>

The NOAA Operational Model Archive and Distribution System (NOMADS) provides distributed, web-service access for real-time and retrospective, format-independent climate and weather model data and related datasets.

NOAA NCDDC, with partners including National Weather Service National Centers for Environmental Prediction (NCEP) and the Northern Gulf Institute, has created this NOMADS node for ocean-model access, called OceanNOMADS. This site provides retrospective access to long timeseries of output from mature ocean modeling and

Point of Contact
oceanNOMADS@noaa.gov
866.732.2382



Access by Data Type for FY2015

Data Type	Volume (TB)	Requests
CFSR	539.13	72,052,889
GFS	286.18	53,725,780
NARR	104.09	51,748,222
NAM	63.64	3,631,991
GDAS	24.98	347,097
RUC / RAP	21.65	4,379,308
NDFD / NDGD	18.06	12,655,678
GENS	2.35	27,198,408
NCEP / DOE Reanalysis-2	0.39	29,394
Tape Orders / Other	264.55	22,297,608
Total	1325.02 TB	248,066,375



Access Tools for Users

- Pare down large file sizes of high resolution data and products- and provide flexible inter-operable access.
- (re-) Group different data sets to create needed products – such as initialization files for model development, analysis, or by forecast projection.
- Subset and aggregate model data:
 - in parameter space
 - in physical space
 - in temporal space

National Academies and NOAA CIO Recommendations

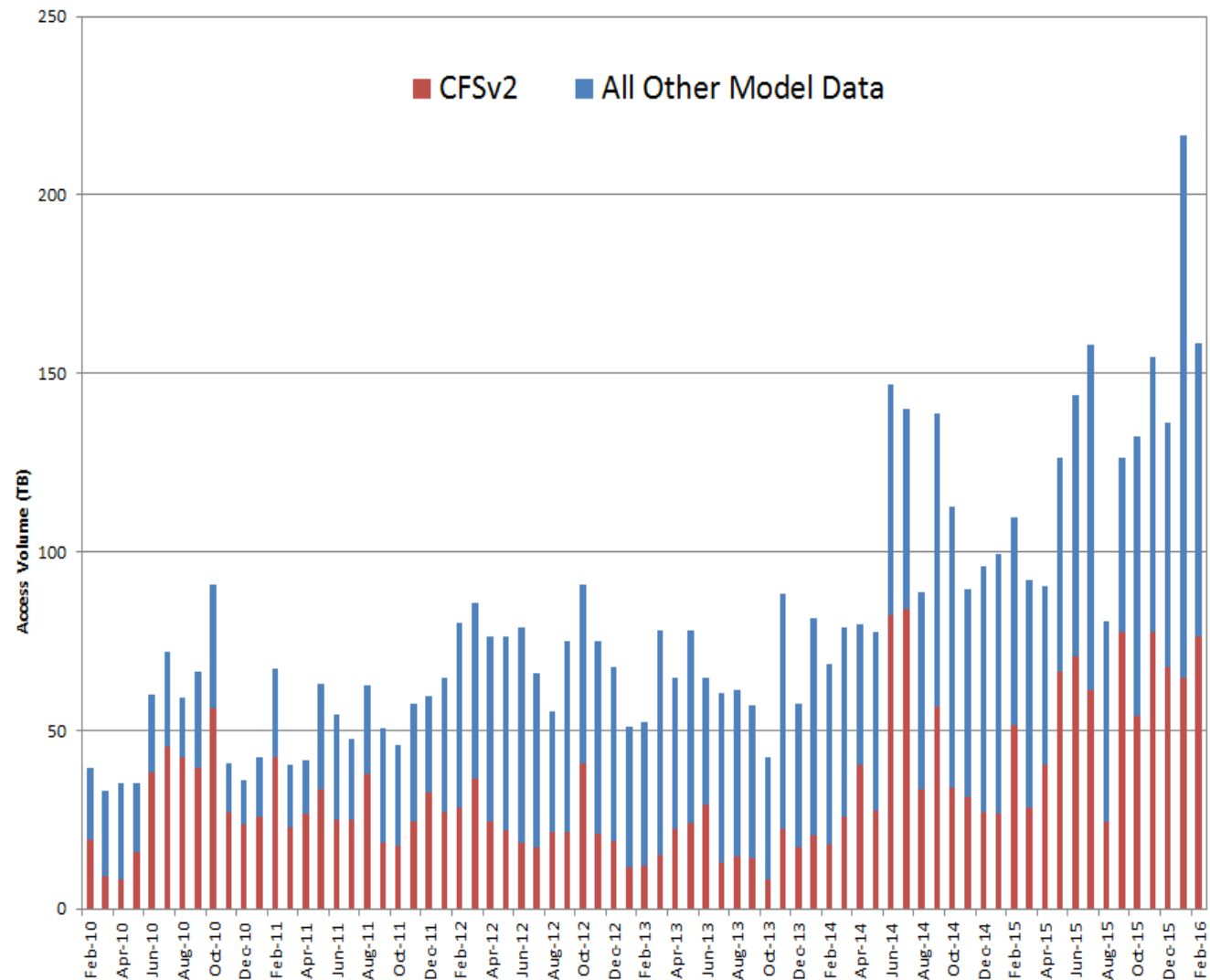
- National Academies of Sciences, NRC, Board of Atmospheric Sciences and Climate (BADC):

“Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts” ***The NOAA National Operational Model Archive and Distribution System (NOMADS)*** should be maintained and extended to include (a) long-term archives of global and regional ensemble forecasting systems and their native resolution, and (b) re-forecast datasets to facilitate post-processing”

- NOAA Office of the CIO Strategic Information Technology Plan 2010-2017, March 2010 IT Strategic Objectives (one of nine):

*“Continue development/implementation and maintenance of operational archive and user access capabilities for the **National Climate Model Portal (NCMP)** to provide access to the next generation of NOAA climate and weather models and analyses products.”*

CFSv2 Usage Statistics



FYI: NOAA Data Alliance

- Public-Private Partnerships are now being advanced under the NOAA Big Data Alliance to leverage private sector resources to store NOAA's model data:
 - Amazon Web Services
 - Google Cloud Platform
 - IBM Microsoft
 - Open Cloud Consortium
- See: <https://data-alliance.noaa.gov/>

GEO Strategic Plan

- **Vison for GEO**

- to envision a future wherein decisions and actions for the benefit of humankind are informed by coordinated, comprehensive and sustained Earth observations.

- **GEO's Mission**

- to connect the demand of decision-makers with the supply of data and information
- to unlock the power of Earth observations

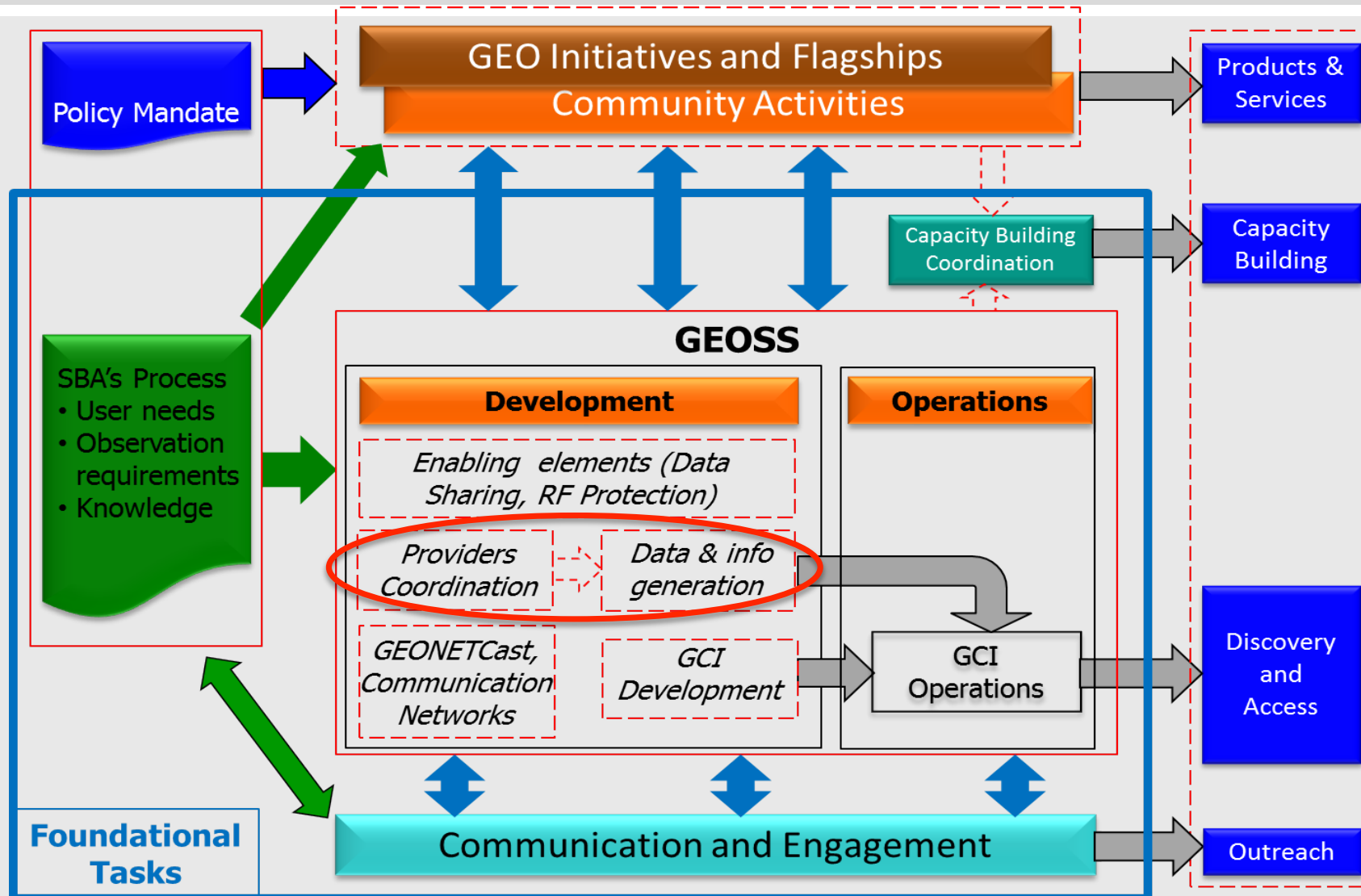
- **GEO's Value**

Convening Power

- to assemble and coordinate expertise from across different disciplines and communities



GEOSS and Workflows



GEOSS Implementation Requires:

Data Sharing Principles

- **Full and Open Exchange of Data**
- **Data and Products at Minimum Time Delay and at Minimum Cost**
- **Free of Charge or Cost of Reproduction**

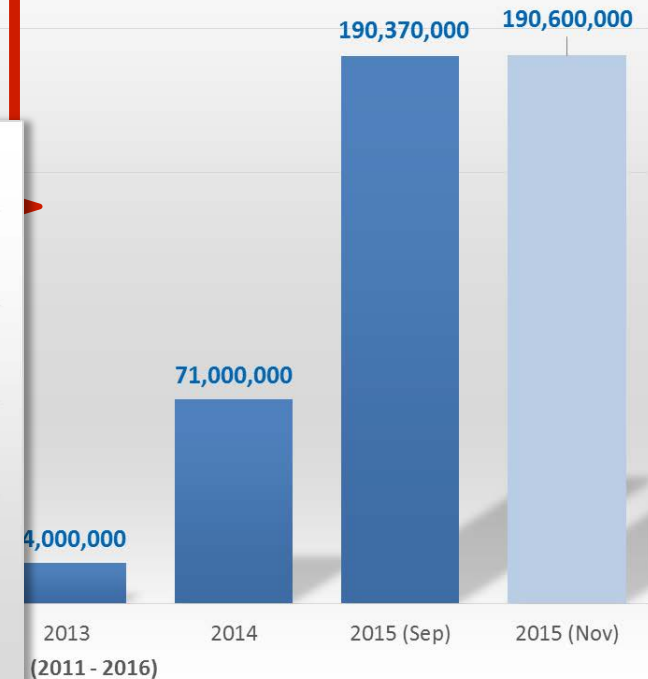
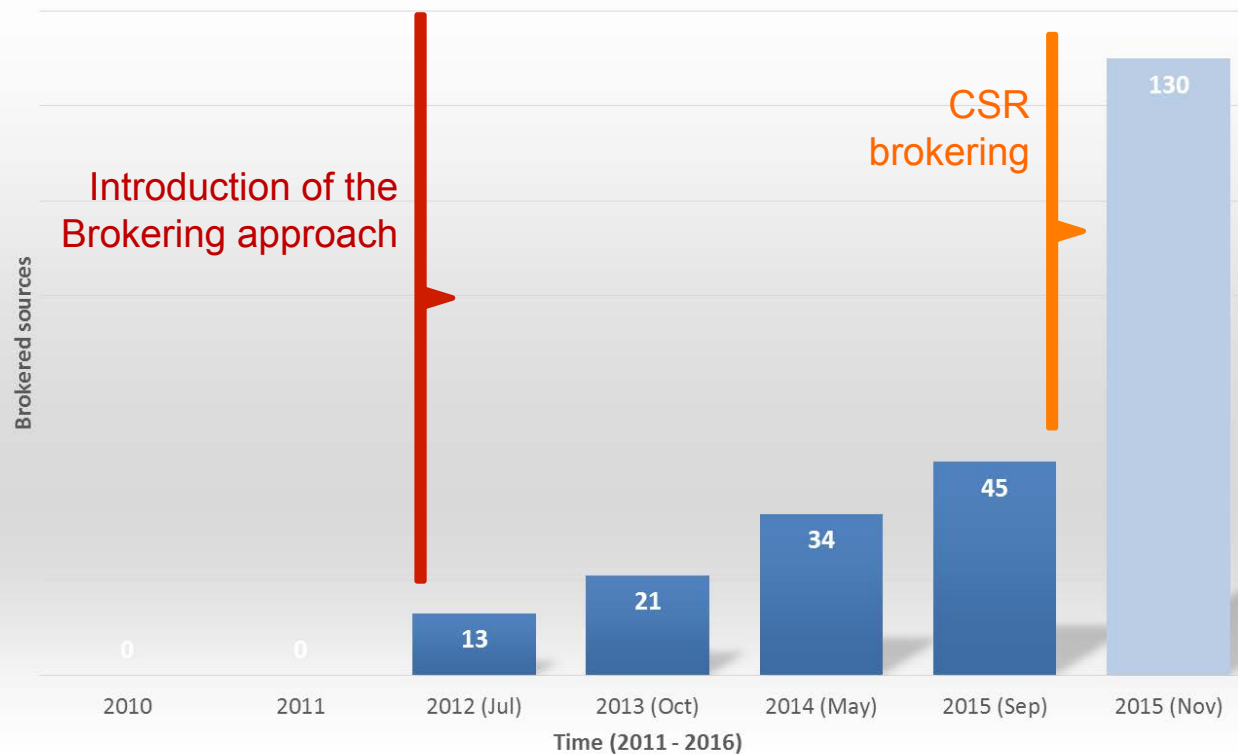


GCI Providers Evolution and Statistics

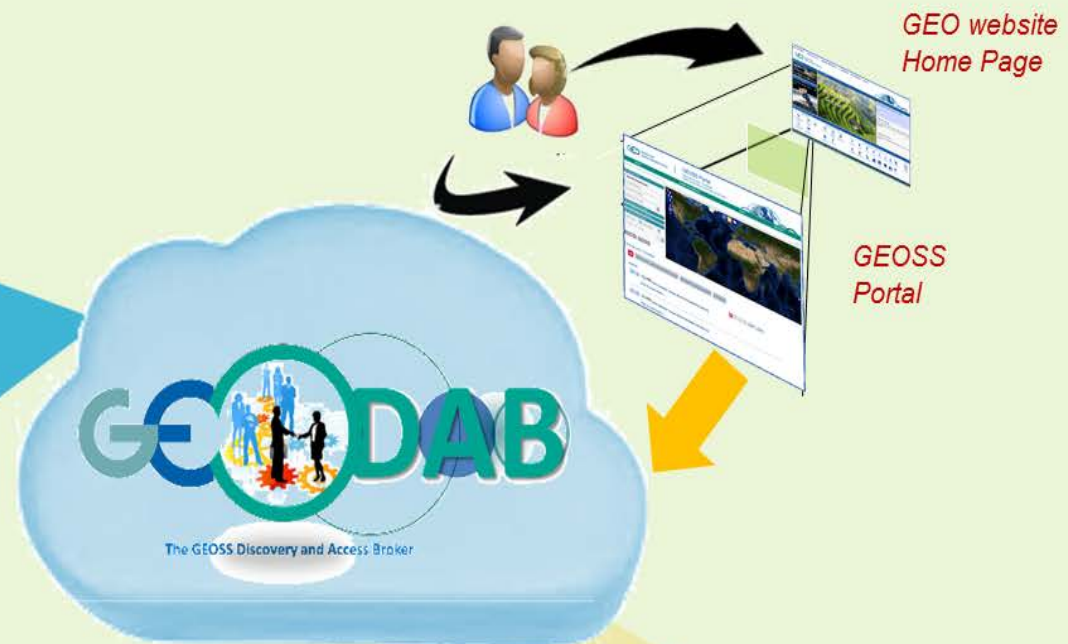
Discoverable and potentially Accessible Elements

Introduction of the

Number of Brokered Enterprise Systems



GEOSS Information System



2016-2015 Strategic Plan

- **GEO' Strengths**

- a unique, multidisciplinary initiative that occupies an upstream coordination position
- a flexible and agile forum
- a facilitator of policy-level

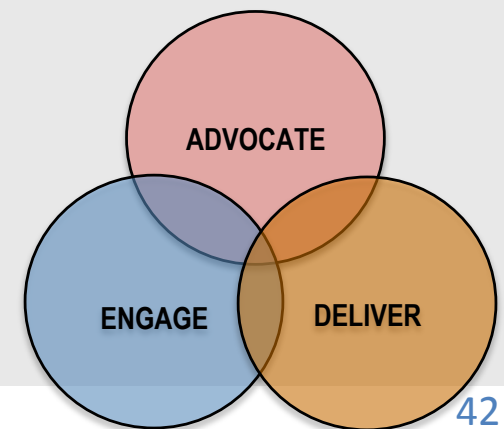
- **GEO's Scope**

end-to-end process:

- identifying data needs
- ensuring the availability of data
- transforming the information into knowledge

- **GEO's Strategic Objectives**

- to advocate
- to engage
- to deliver



GEO Task CA-03 “Access to Climate Data in GEOSS”

Overarching Activities and Outcomes

- An effective international federated and distributed data service requires coordination of data infrastructure and data management principles extending beyond organizational boundaries of any individual center.
- The fundamental data management issue that CA-03 will address is how GEO institutions can organize their distributed data resources into a cohesive presence, to allow the users to make better, informed decisions about how nature will impact their future, either in their life or in their business decisions.



GEO Task CA-03

- CA-03 “Access to Climate Data in GEOSS”
- Primary focus of this activity is a cluster of 3 main themes:
 - 1) data discoverability, 2) archive access, and 3) dissemination.
 - It is envisioned this Task will assist users of all levels of expertise to find and discover climate and weather information including access tools and visualization capabilities.
 - These objectives will be advanced by leveraging and coordinating with selected national and international data, modeling and information access groups including WCRP, GCOS, WMO GFCS.



GEO Task CA-03 Next Steps 1/2

- Leverage the IPCC WCRP CMIP infrastructure for access to climate model outputs under the ESGF international collaboration. Advance the NCEI-NOMADS-CICS-NC ESGF Pilot;
- Access to associated observational data sets under the ESGF/obs4MIPs/ana4MIPs activities will be championed within a wider (beyond climate research) community;
- The U.S. NOMADS model data access system, already a GEO contributor will advance CA-3 goal and objectives;
- Promote the wider development and use of ESGF for climate simulations at all spatial and temporal scales and Earth System domains, including regional downscaling (CORDEX), seasonal and decadal predictions and WCRP core projects model development and inter-comparison initiatives;

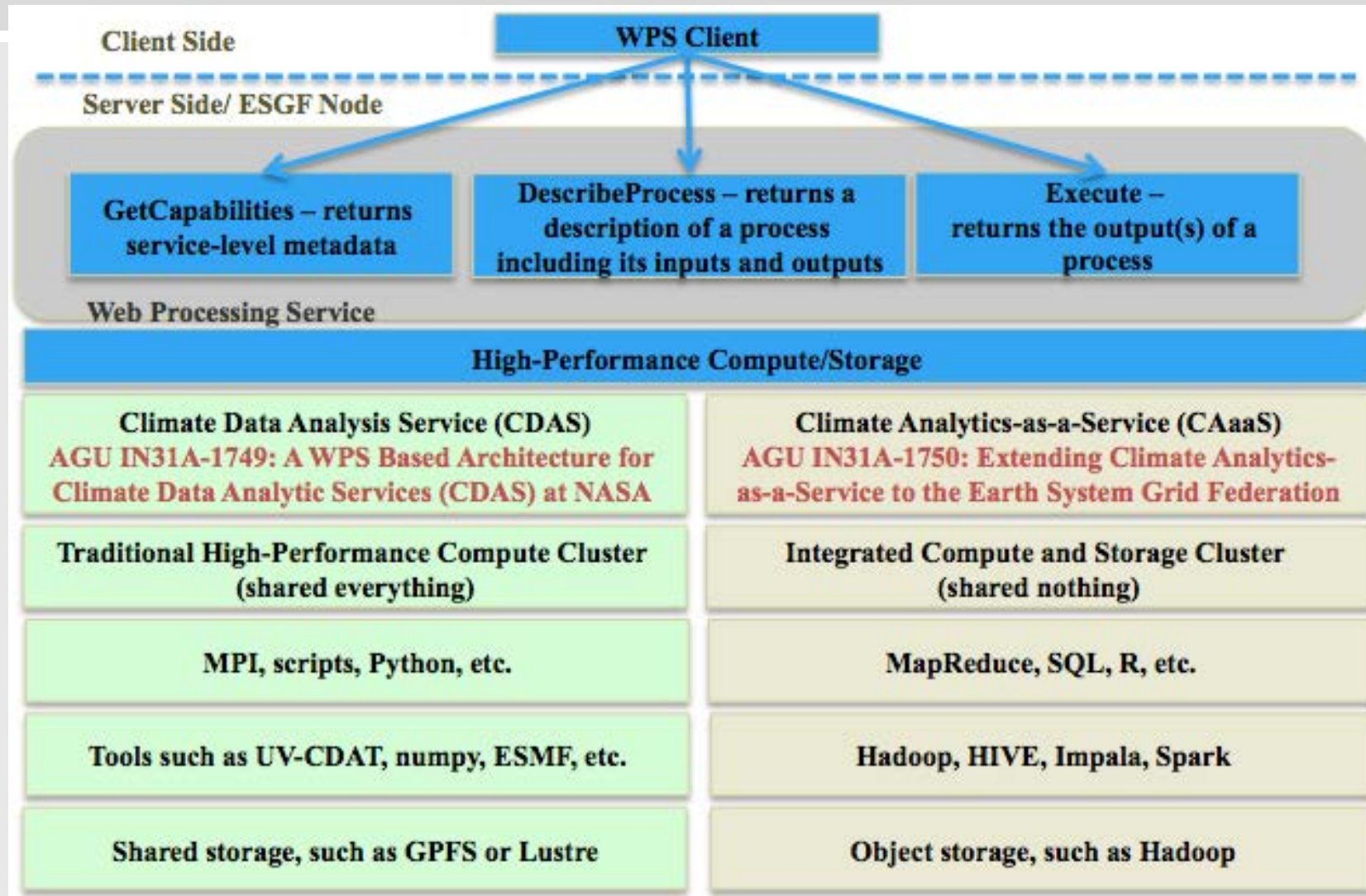


GEO Task CA-03 Next Steps 2/2

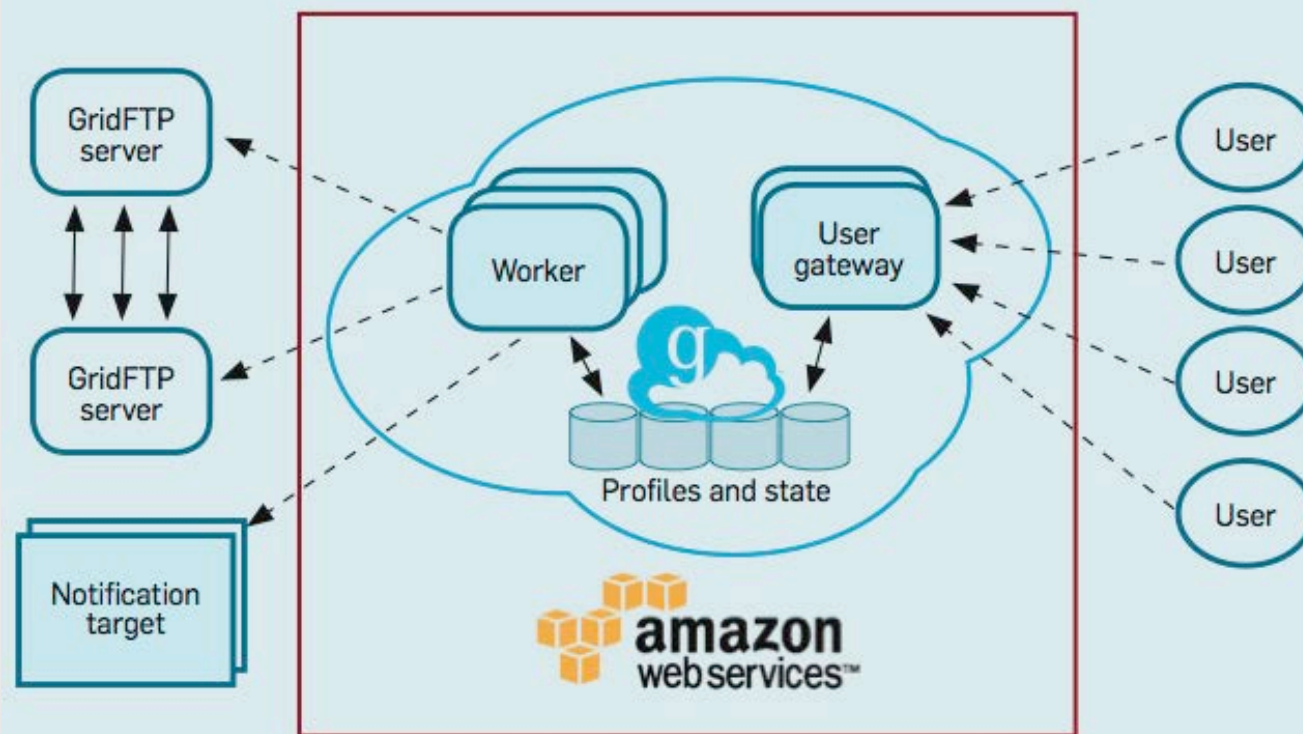
- Collaboration with the Decadal Forecast Exchange data and links with contributions from the Copernicus Climate Change Services C3S Project;
- Promote product development and collaborations within the geo-science communities (ocean, weather, and climate) to foster interdisciplinary research to study multiple earth systems using collections of distributed data under a sustainable GEOSS Common Infrastructure (GCI);
- Advance GEO collaborations and linkages to NASA's CREATE (reanalysis clearinghouse); and Reanalysis.org



Two reference back-end



Globus Online Architecture



Cloud-based services for big-data fire-and-forget file transfer

Source: [*Software as a Service for Data Scientists; Allen et. al; Argonne Nat'l Lab; 2011*](#)



ON-DEMAND STREAMING

of massive climate simulation ensembles

Cameron Christensen*
Giorgio Scorzelli*

Peer-Timo Bremer^o
Valerio Pascucci*



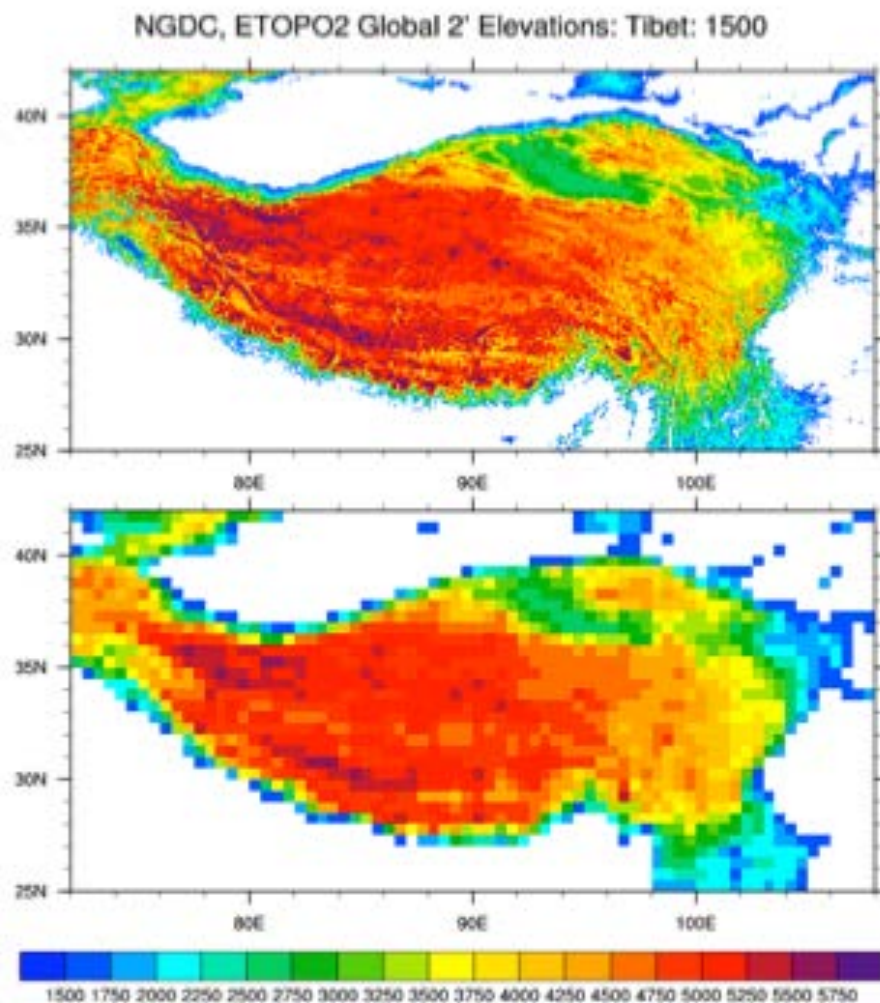


Visualization streaming

ANALYSIS AND VISUALIZATION

Click to add text

- Retrieve datasets to common location
- Regrid data to common domain
- Process using out-of-core or parallel supercomputer
- Downsample large domains or visualize using distributed parallel cluster application
- Store results
- Share

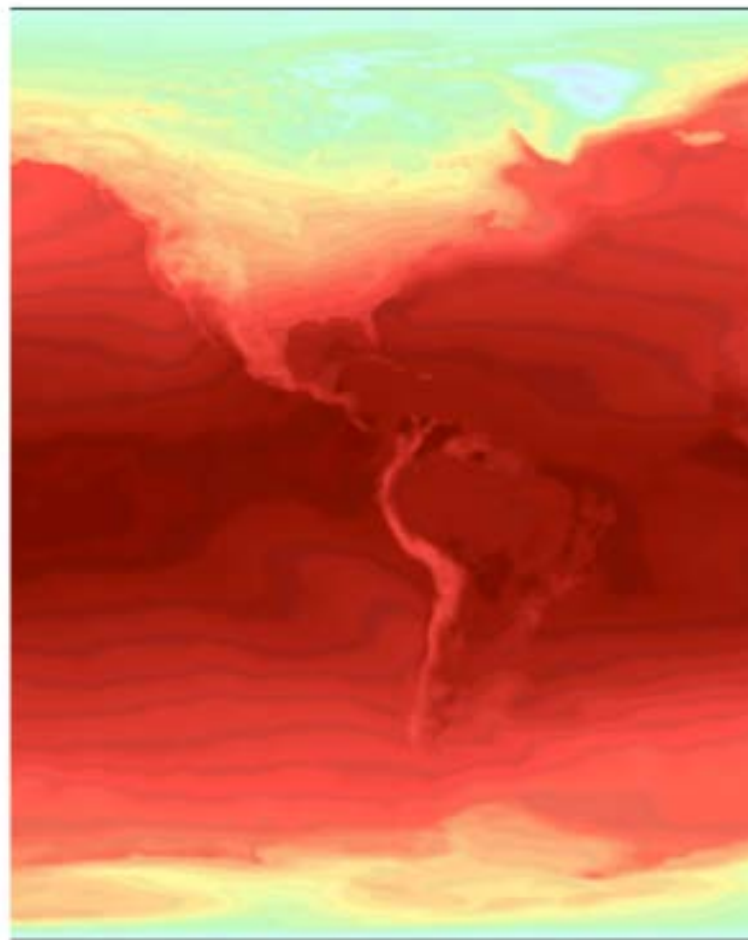




Stream multi-resolution

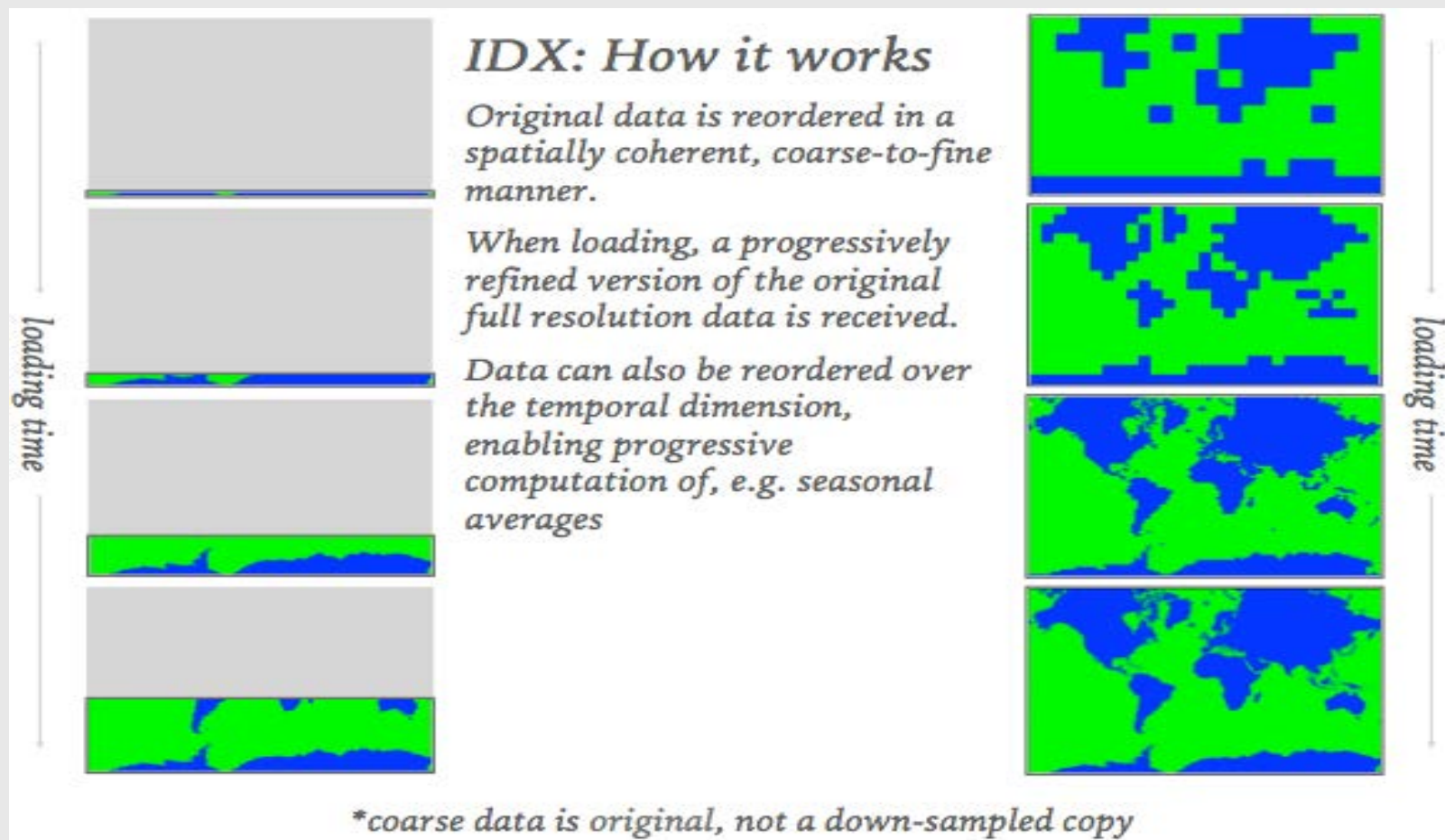
STREAMING MULTIREOLUTION

- Interactive view-dependent data loading ➡ permits flexible exploration outside fixed regions
- Fast coarse resolution results for visualization and analysis
- High resolution results through user-directed refinement
- Low requirements
 - low bandwidth: ➡ download less
 - low storage and computation ➡ process and cache only what you download
 - fast
- *you already use it every day!*





IDX format





Accelerated Climate Modeling for Energy (ACME) end-to-end workflow

