Scaling the Earth System Grid to 100Gbps Networks

Mehmet Balman, Alex Sim

Scientific Data Management Research Group
Computational Research Division
Lawrence Berkeley National Laboratory

100Gbps Network and Climate Community

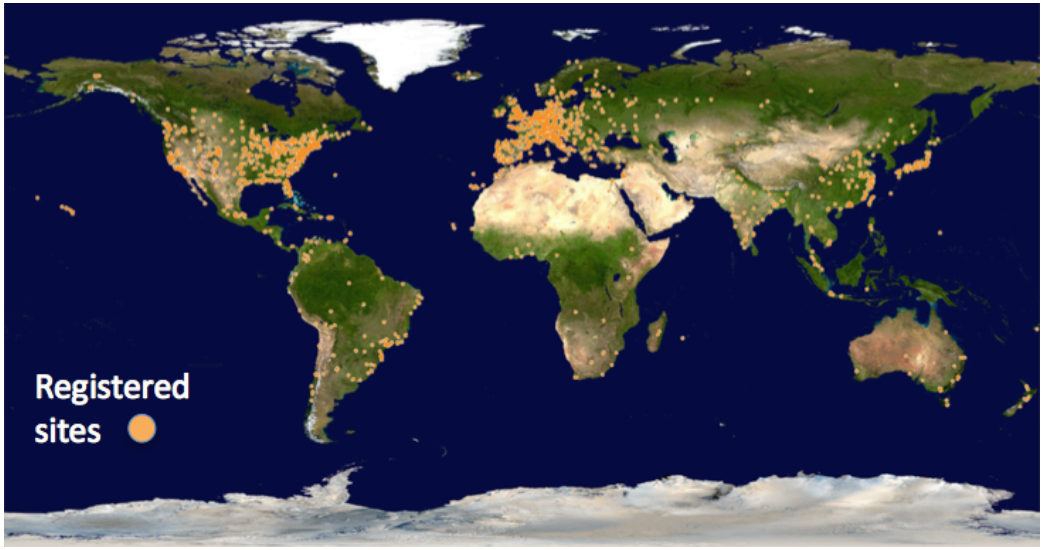
- **High performance network for climate community**
  - **Data replication around the world**
    - **"Replica Core Archive" – The Coupled Model Intercomparison Project, Phase 5 (CMIP-5) used for the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) is estimated to 1.5-2PB.**
    - **Climate model data is projected to exceed hundreds of Exabytes by 2020 (BES Science Network Requirements Workshop, 2007)**
  - **Data analysis**
    - **Climate analysis requires dataset**
    - **Large amount of data movement and management is needed**
- **Climate100**
  - **Research and integration effort for 100Gbps network from data intensive applications point of view**

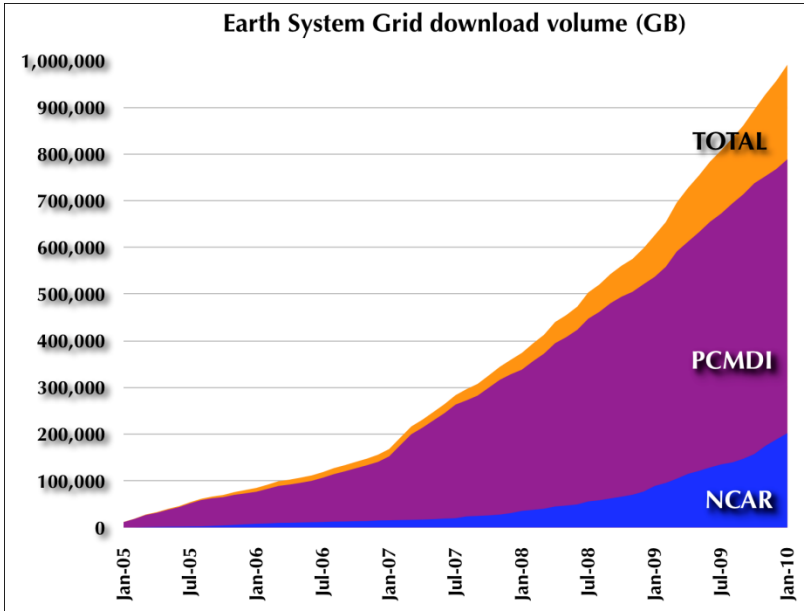
ESG current statistics

- **IPCC AR4 CMIP-3 LLNL ESG portal**
  - 35 TB of data at one location
    - 83,337 files, model data from 13 countries
    - Generated by a modeling campaign coordinated by the Intergovernmental Panel on Climate Change (IPCC)
    - Over 600 scientific peer-review publications
- **CCSM NCAR ESG portal**
  - 237 TB of data at four locations (NCAR, LBNL, ORNL, LANL) : 965,551 files
    - Includes the past 8 years of joint DOE/NSF climate modeling experiments
- **Serving data to the community**
  - Coupled Model Intercomparison Project, Phase 3 (CMIP-3)
  - Community Climate System Model (CCSM)
  - Parallel Climate Model (PCM)
  - Parallel Ocean Program (POP)
  - The North American Regional Climate Change Assessment Program (NARCCAP)
  - Cloud Feedback Model Intercomparison Project (CFMIP)
  - Carbon-Land Model Intercomparison Project (C-LAMP)

ESG current download statistics

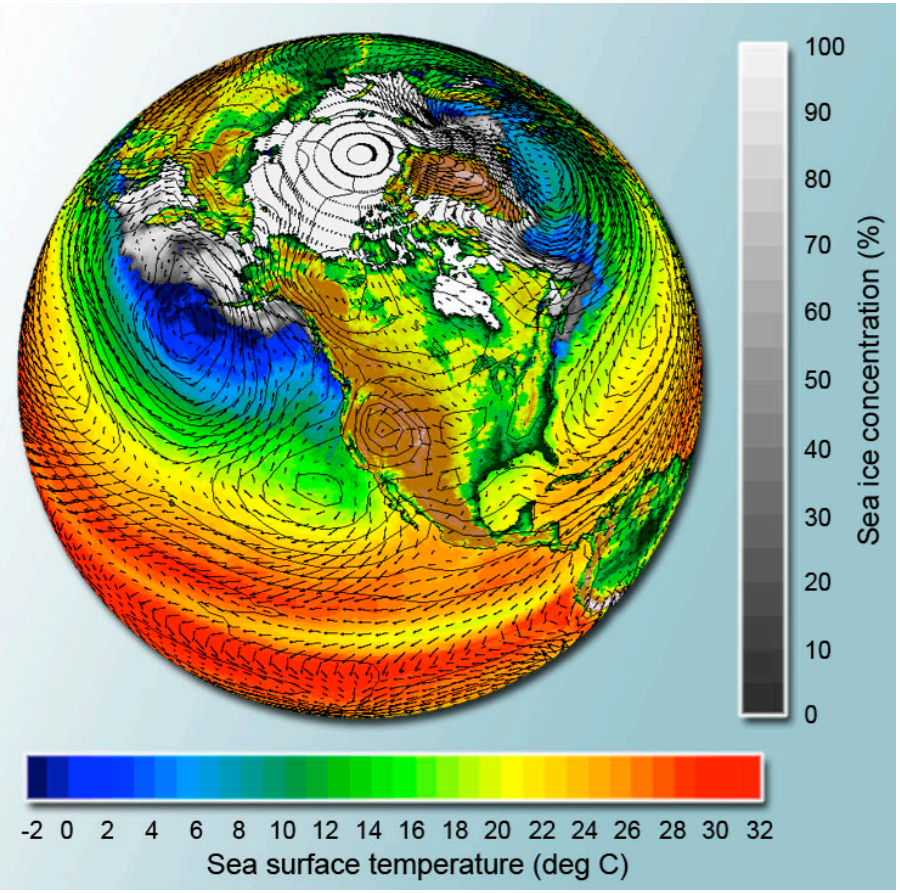
- **Geographic distribution of the users that downloaded data from ESG web portals**
  - Over 2,700 sites
  - 120 countries
  - 25,000+ users
  - Over 1 PB downloaded





The Growing Importance of Climate Simulation Data

- **Broad investments in climate change research**
  - Development of climate models
  - Climate change simulation
  - Model intercomparisons
  - Observational programs
- **Climate change research is increasingly data-intensive**
  - Analysis and intercomparison of simulation and observations from many sources
  - Data used by model developers, impacts analysts, policymakers

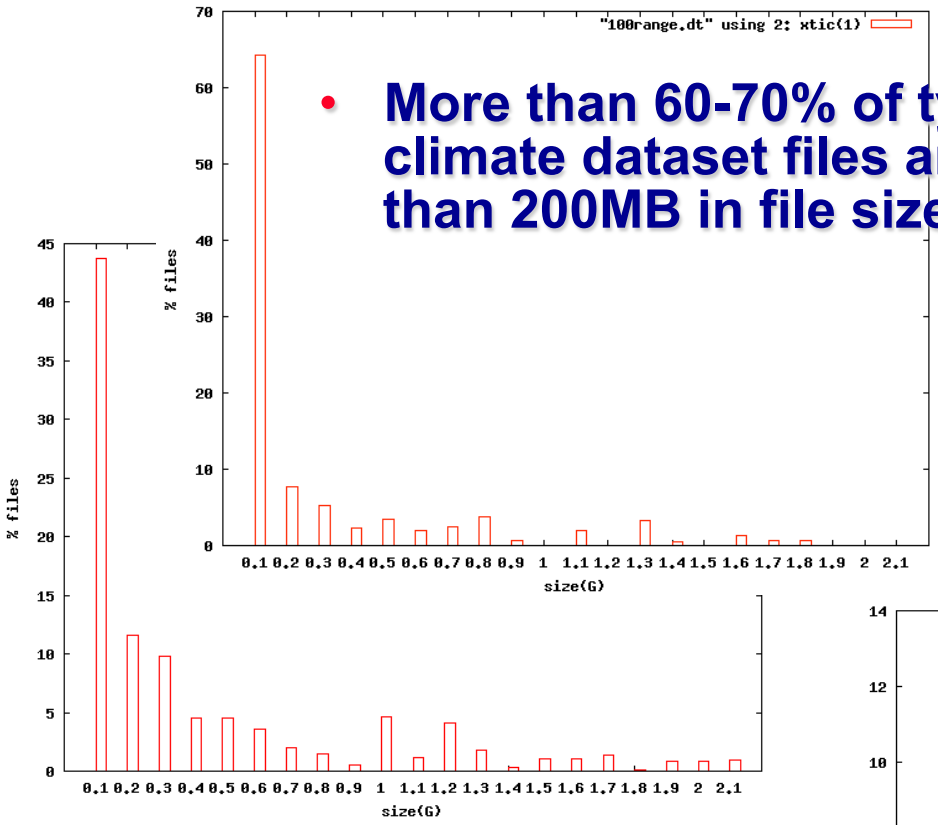


Results from the Parallel Climate Model (PCM) depicting wind vectors, surface pressure, sea surface temperature, and sea ice concentration. Prepared from data published in the ESG using the FERRET analysis tool by Gary Strand, NCAR.

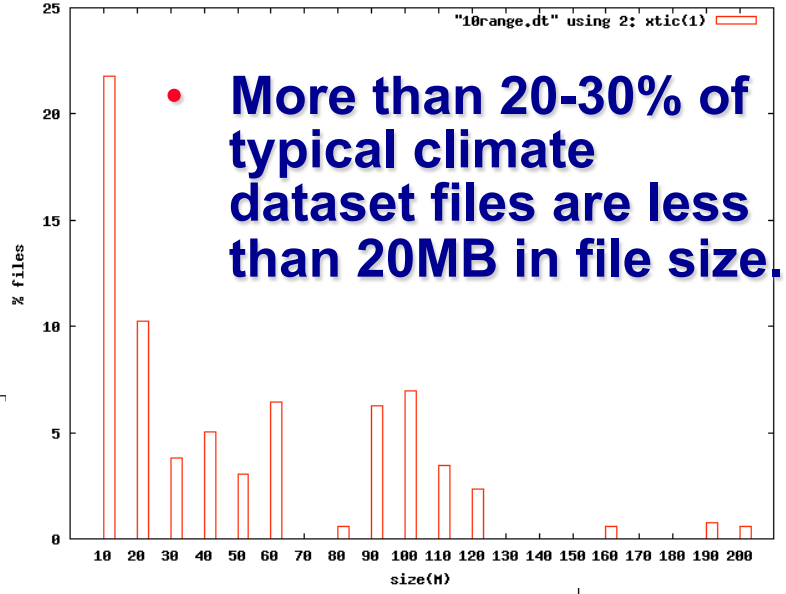
The Growing Size of Climate Simulation Data

- **Early 1990's (e.g., AMIP1, PMIP, CMIP1)**
  - modest collection of monthly mean 2D files: ~1 GB
- **Late 1990's (e.g., AMIP2)**
  - large collection of monthly mean and 6-hourly 2D and 3D fields: ~500 GB
- **In 2000's (e.g., IPCC/CMIP3)**
  - fairly comprehensive output from both ocean and atmospheric components; monthly, daily, and 3 hourly: ~35 TB
- **In 2011:**
  - The IPCC 5th Assessment Report (AR5) in 2011: expected 5 to 15 PB
  - The Climate Science Computational End Station (CCES) project at ORNL: expected around 3 PB
  - The North American Regional Climate Change Assessment Program (NARCCAP): expected around 1 PB
  - The Cloud Feedback Model Intercomparison Project (CFMIP) archives: expected to be .3 PB
- **CMIP5 is being defined now, available info neither complete nor final**
  - Current estimates… 1.5 to 2 PB of "replica core archive" (RCA) results
    - In CMIP5, the RCA is expected to be 20% to 30% of total volume of data produced
- **Climate model data is projected to exceed hundreds of Exabytes by 2020**
  - BES Science Network Requirements Workshop, 2007

Distribution of Climate Files – Characteristics of CMIP-3 datasets

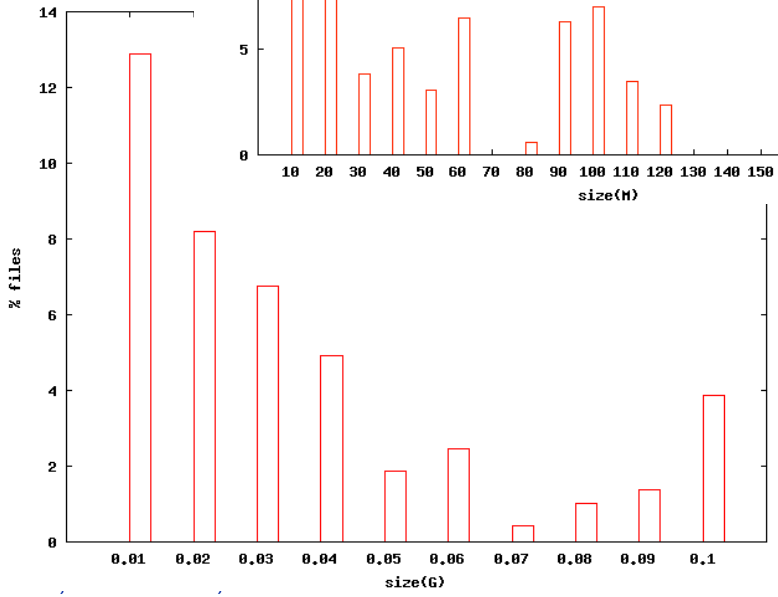


- More than 60-70% of typical climate dataset files are less than 200MB in file size.

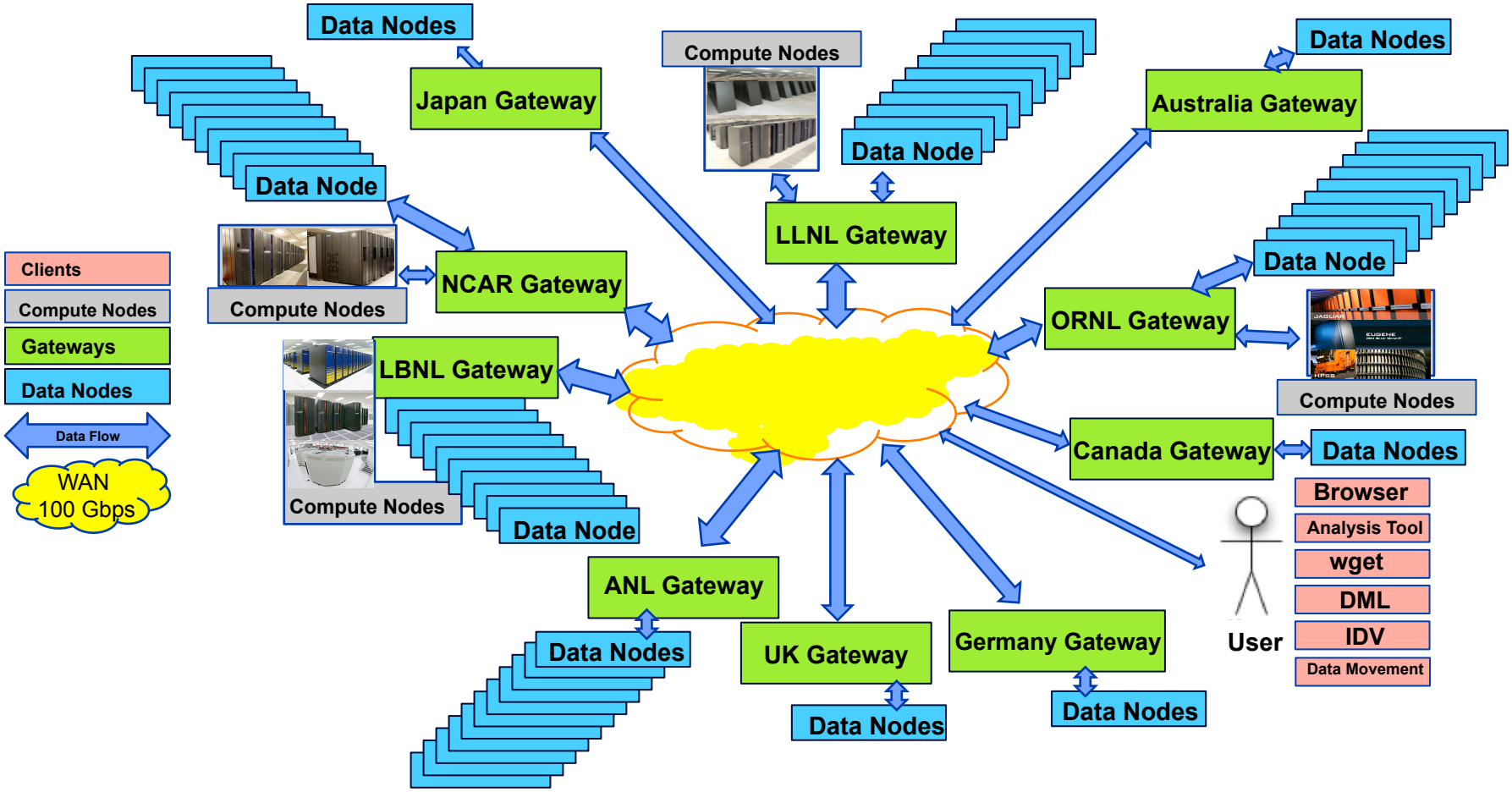


- More than 20-30% of typical climate dataset files are less than 20MB in file size.

- Many files are still larger than 2GB.
- It's expected to be as large as 20+GB per file for CMIP-5 datasets in the near future.



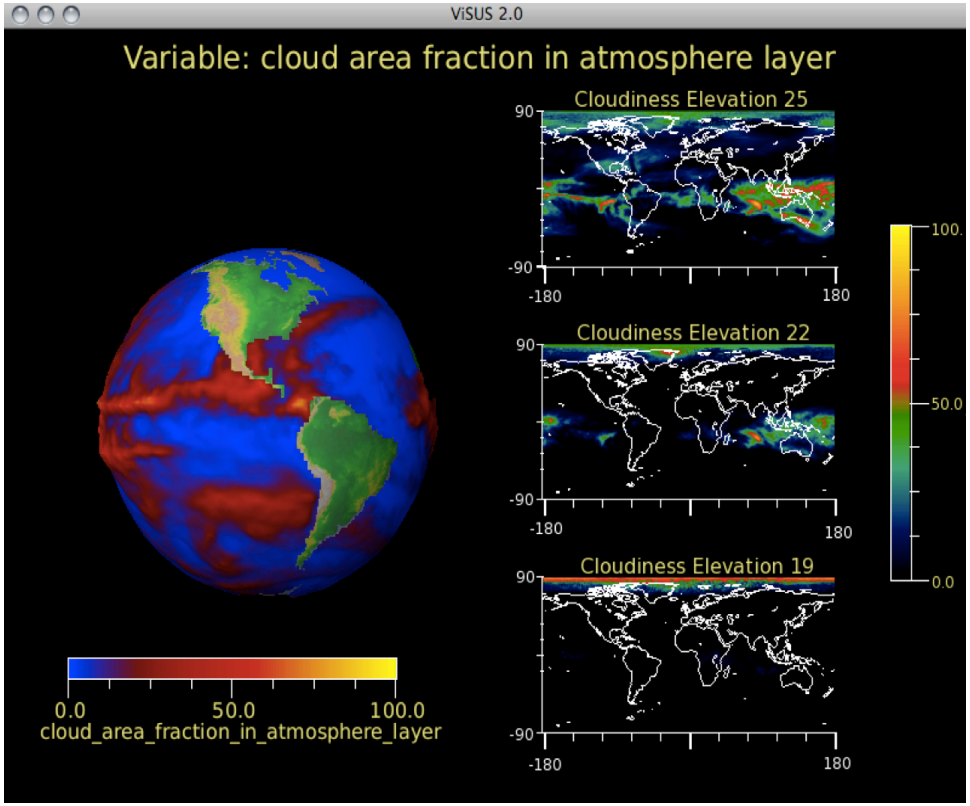
Network Challenges in ESG

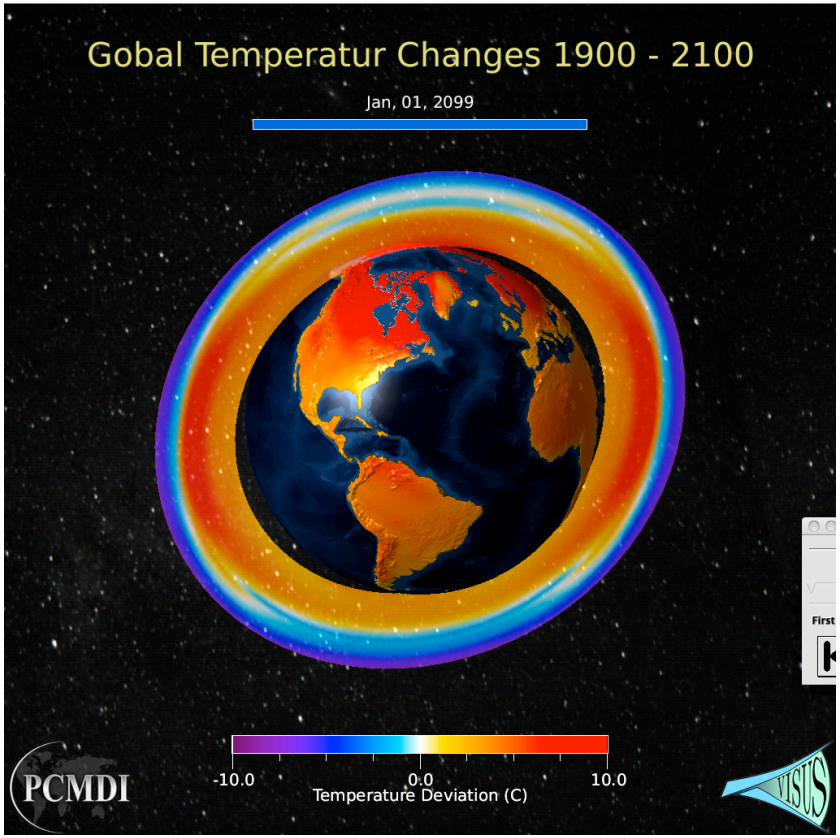


- ESG Federation plans to replicate vast amounts of data to and from sites
- ESG Federation plans to utilize current ESnet's 10 Gbps network for its federated architecture, and move onto 100Gbps when ready

As the results...

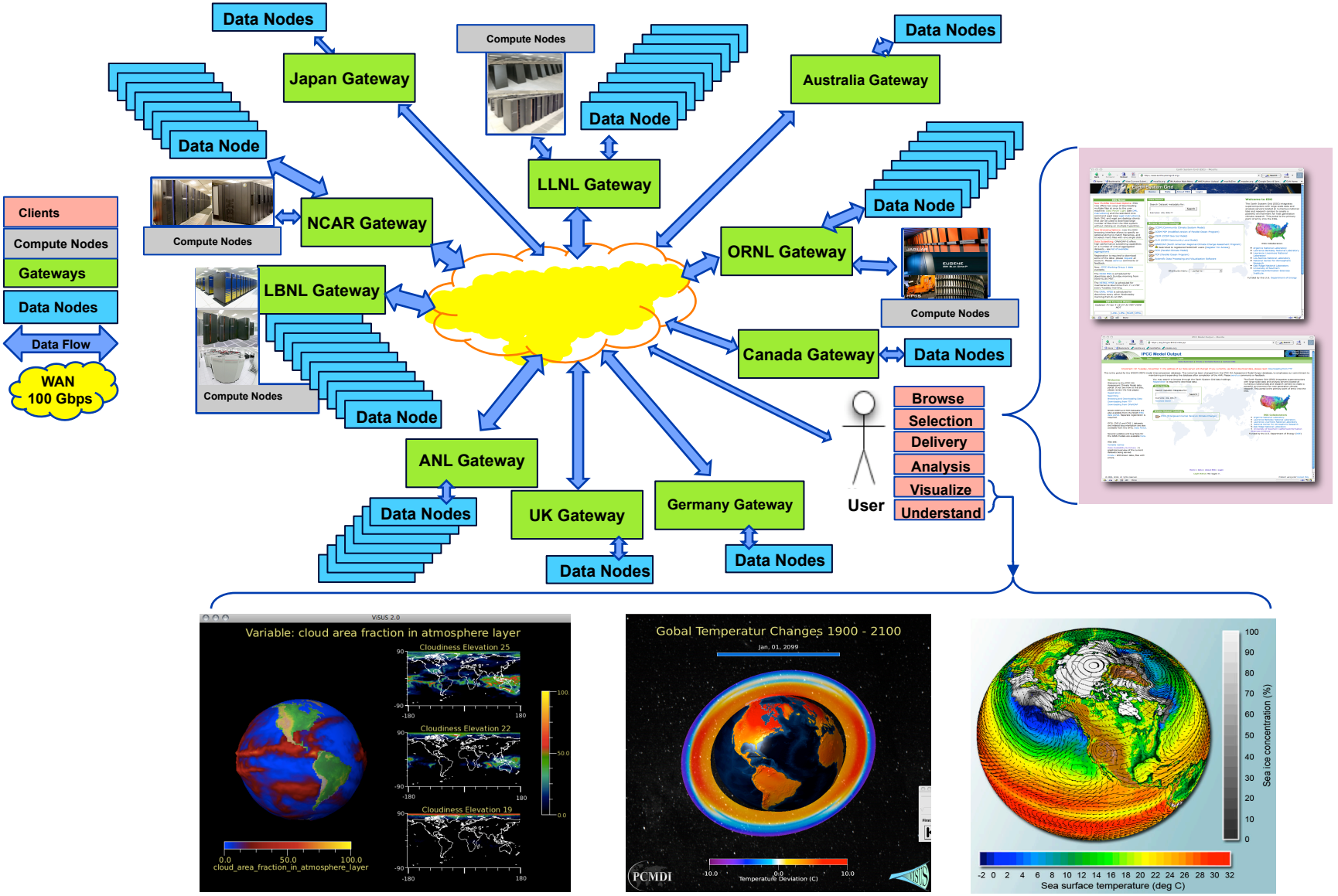
- **Enable new capabilities for analysis of data and visual exploration**
  - Visualization of uncertainty and ensemble data
  - Exploration of climate modeling data with ViSUS and CDAT
- **Help scientists understand long-term climate impact**





Scaling climate community to 100 Gbps networks

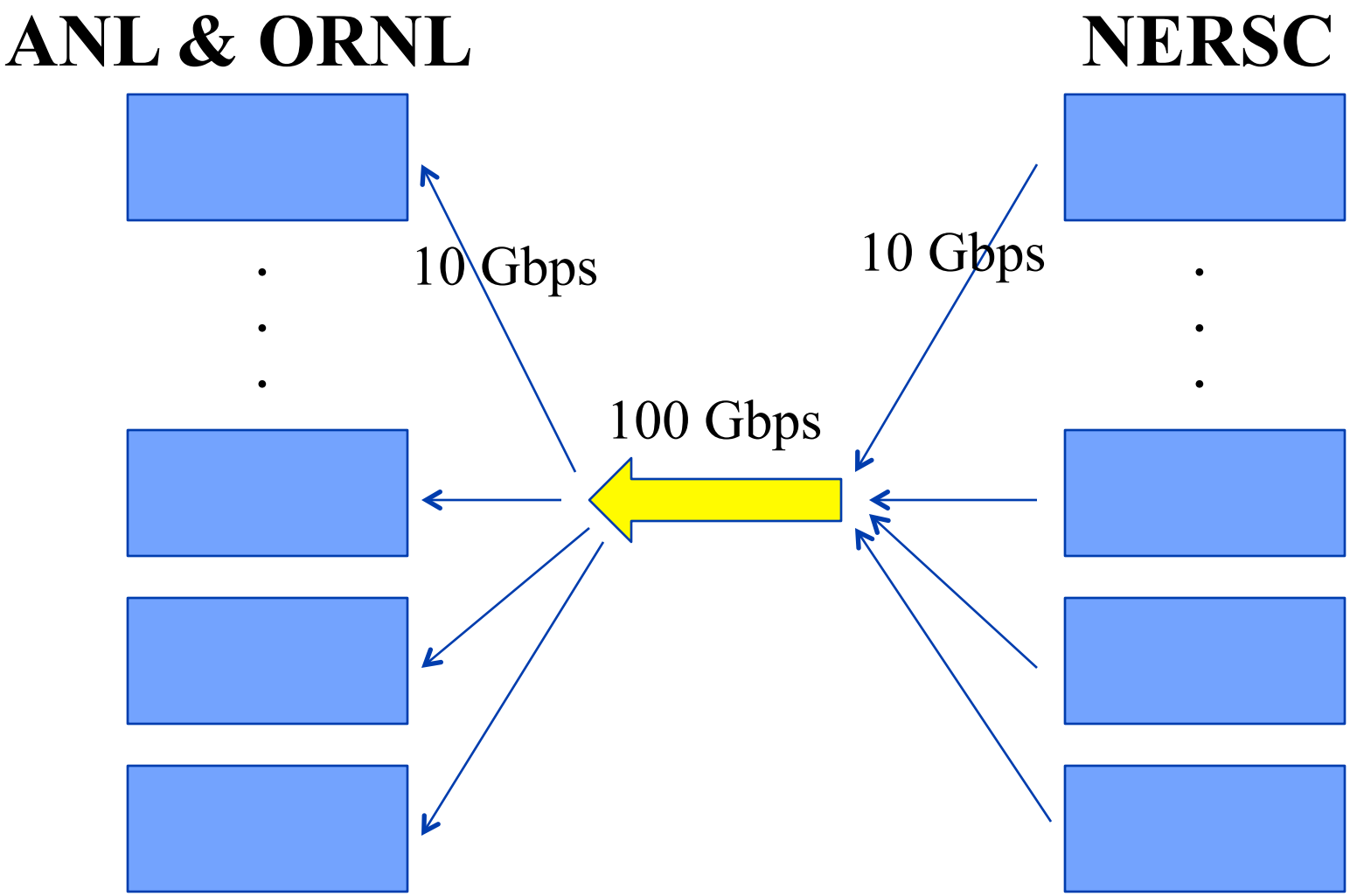
Earth System Grid



ANI 100 Gbps Testbed

- **Testbed**
  - LBNL/NERSC: 20 nodes
  - ANL/ALCF: 15 nodes
  - ORNL/OLCF: 15 nodes
  - Each connection with a 10 Gbps interface
- **Climate demo**
  moving datasets of IPCC AR4 CMIP3
  - From NERSC to ANL over 100Gbps
    - Disk to memory
  - From NERSC to ORNL over 100Gbps
    - Disk to disk

Data flow over 100Gbps network

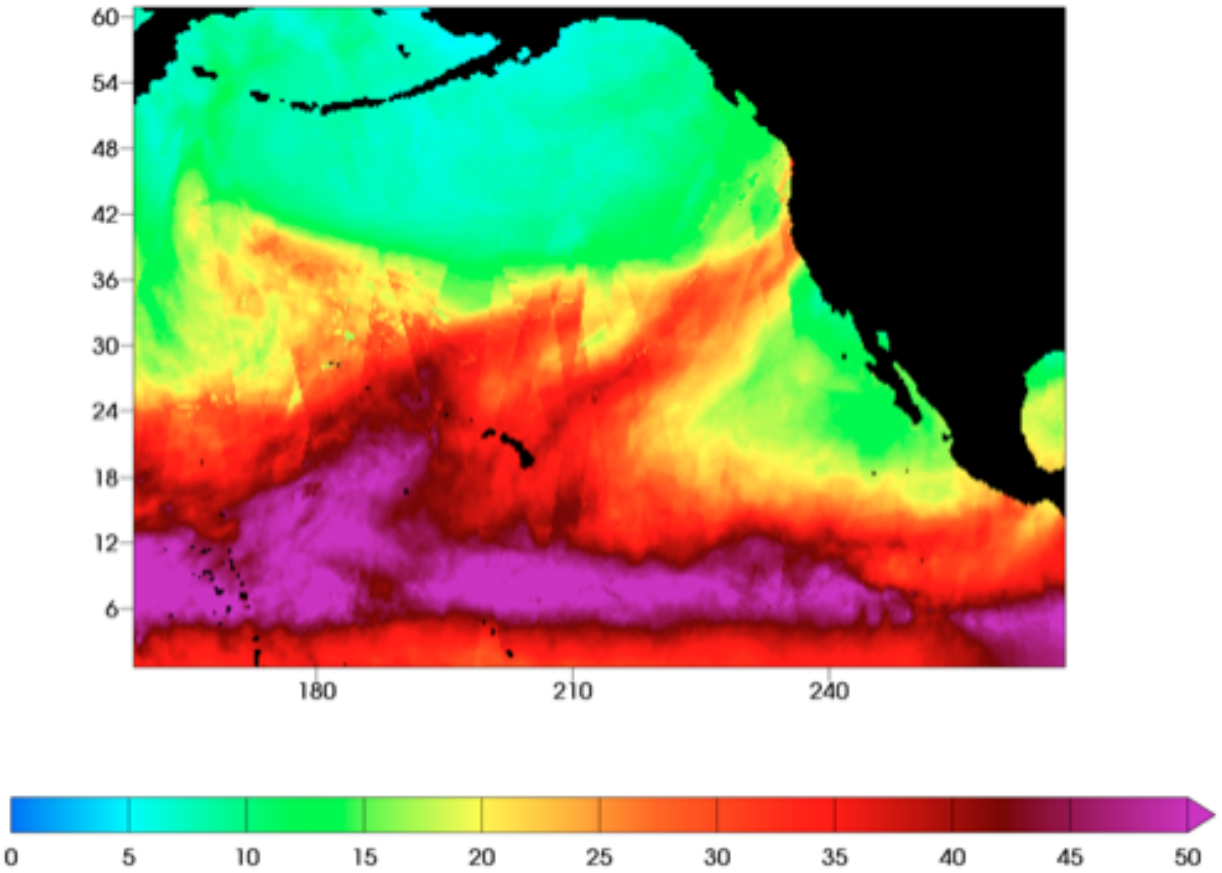


Showcase

- Use of maximum available bandwidth for the movement of climate data over 100Gbps network
- Possible challenges
  - Using existing tools in 100 Gbps networks
    - GridFTP for TCP
    - RFTP for RDMA
  - Irregular file size distribution in each dataset
    - Protocol overhead
  - Performance problem and scalability issues
    - Management and tuning of multiple hosts
    - Number of involving host systems
    - Multiple streams for increased utilization
    - Performance monitoring in host systems
    - System bottleneck in end-to-end transfers
    - Memory overhead / CPU usage

Sample Analysis on Large Climate Dataset for the Demo

Detecting Atmospheric Rivers



Source: S. Byna et al., Detecting Atmospheric Rivers in Large Climate Datasets. SC'11 Workshop. 2011.