Advancing Capabilities for Climate Data Access and Analysis

Don Middleton
(on behalf of a lot of people and projects!)
National Center for Atmospheric Research
Scientific Computing Division
Section Head, Visualization & Enabling Technologies

NCAR
1. Community technology development efforts are making it possible to deliver new web-based capabilities rather easily…
The Community Data Portal

A new environment and suite of services aimed at enabling both the consumers as well as the suppliers of important community datasets.

- Shared hardware/software infrastructure that facilitates broad access and low cost-of-entry for provision
- Multiple access methods for shared data objects
- A flexible multi-technology, multi-service approach
- A generalized portal for scientific data
ACACIA Regional Climate Data Access System

CSM1.2 (Business as usual) served by SCD/NCAR
Clearsky net longwave flux at surface

Select view
- xy (lat/lon) slice

Select
- single variable
- comparison

Get Data
- Full Region

Select time
- 01 Jan 1980
- 01-Jan-1980

Select product
- Shaded plot (GIF)
- 800x600 window
VEMAP

The Vegetation/Ecosystem Modeling and Analysis Project Data Server

Select view: xy (lat/lon) slice

Select: single variable

Select product: Shaded plot (GIF)

Select time: 01 Jul 1985 01-Jul-1985

BBGCC1SuCVA404i served by Experiment C1SuC
Annual Mean Total Carbon Storage

Select a region:

48.2917 N
124.25 W
67.458 W
25.25 N

Zoom In
Zoom Out
**Atmospheric Profiler**

**Profiler**

National Center for Atmospheric Research

- **Select First Case**
  - csu310
  - ccm3527
  - L26
  - r1up
  - ccm3511
  - rasL26
  - ras03
  - zhang
  - csuL26
  - sld031
  - trig6
  - vdt1
  - e0220

- **Select Comparison**
  - csu310
  - ccm3527
  - L26
  - r1up
  - ccm3911
  - rasL26
  - ras03
  - zhang
  - csuL26
  - sld031
  - trig6
  - vdt1
  - e0220

- **Choose Station**
  - **Most Common:**
    - Chuuk Island
    - Yap Island
    - Easter Island
    - Gilbert Islands
    - Heard Island
    - Marshall Islands
    - Midway Island
    - Port Moresby, PNG
    - Raou Island
    - Samo Island
    - Singapore
    - Sydney
    - Tokyo
  - **All others:**
    - Alaska, Gulf of
    - Alaska, Western
    - Antarctic, Ross Ice Shelf
    - Ascension Island
    - Atlantic, North

- **Select Month**
- **Variable**
  - Temperature
  - Specific Humidity
  - Geopotential Height
  - Dry Static Energy
  - Moist Static Energy

- **All 4 Panel Plot**
- **Or Choose:**
  - January
  - April
2. The Future Poses Formidable Challenges
The Earth System Grid (ESG)

http://www.earthsystemgrid.org

- Funded by the U.S. DOE Scientific Discovery through Advanced Computing (SciDAC), this program seeks a new paradigm in the climate change community evolving from centralized data sharing to distributed data-sharing.

- Enabling geographically distributed teams of researchers to effectively and rapidly acquire knowledge and understanding of massive amounts of climate data holdings.

- Multiple interfaces to ESG will allow researchers to focus on science and not issues with data receipt, format, and data set manipulation.
What is “The Grid”? 

Central Concept: “Coordinated resource sharing and problem-solving in dynamic multi-institutional virtual organizations”

• Analogous to the “power grid”
• A digital “world is my oyster”
• A “megatrend”…
• Foundations for a meta-OS?
The Globus Toolkit™

An Open Source Effort

- Security
- Directory Services (LDAP, futures)
- Resource Management
- High-performance Data Access and Management
- Distributed computation
- Future: Reliable, persistent web services
Corporate Commitments…

- Compaq
- Cray
- Sun
- SGI
- Veridian
- Entropia
- Microsoft
- IBM
- NEC
- Fujitsu
- Hitachi
- Platform Computing
ESG Problem Landscape

- Climate research and CCSM model data
- Terascale distributed data
- Primitive management tools
- Massive movement of data on net (slowly)
- Hard to find what you want
- Analysis can be... *challenging*
ESG: U.S. Collaborations & Development

- ORNL: Climate storage & computational resources
- LANL: Next generation coupled models & computing
- ANL: Computational grids, & grid-based applications
- LBNL: Climate storage facility
- LLNL: Model diagnostics & inter-comparison
- USC/ISI: Computational grids, & grid-based applications
- NCAR: Climate change predication and scenarios
- ORNL: Climate storage & computational resources
ESG Focus Areas

- Grid-wide services supporting authentication, authorization, data discovery, and user specified analysis
- Metadata services supporting remote data browsing, querying, accessing, displaying, etc.
- Filtering services performing intelligent model specific analysis before delivering the results to the user
- Integrate next-generation data analysis and visualization applications (such as ongoing work at LLNL and NCAR), web-based data portals and other thin clients supporting the Distributed Oceanographic Data System (DODS), and collaborative problem-solving environments.
ESG Summary

- ESG is a highly collaborative effort and will allow users to quickly access data storage facilities storing petabytes of raw or processed data in an application independent manner.

- Payoffs of this distributed collaborative infrastructure, would include:
  - Distributed data-sharing
  - Simplified data discovery of climate data
  - Large-scale climate data processing and analysis
  - Increased collaboration among climate research scientists
  - Aid in climate assessments and estimates of future climate variability and trends

- For more information on ESG, visit our websites at:
  - http://www.earthsystemgrid.org
3. There is a Gratifying Level of Communication and Cooperation Underway…
ESG
Relationships

- CCSM
- Other DOE SciDAC Projects
- OPeNDAP, Distributed Oceanographic Data System (DODS)
- THREDDS (Unidata et al)
- NOMADS (NOAA)
- U.K. eScience in Climate and Oceans
- CEOS-grid
- NCAR Community Data Portal (CDP)
Towards a Global Federation of Climate Data Access

New Data Access Environments & Tools

CyberInfrastructure for New Community Data Capabilities

ESG
THREDDS
OPeNDAP
NOMADS
U.K. eScience
The Grid And Globus
NCAR
End