Earth Systems Science OPeNDAP
Compute Server Framework

Steven J. Phipps\textsuperscript{1}
Andrew Woolf\textsuperscript{2}
Nathan L. Bindoff\textsuperscript{1}
Glenn B. Hyland\textsuperscript{1}
Jason L. Roberts\textsuperscript{1}

\textsuperscript{1}Tasmanian Partnership for Advanced Computing, Australia
\textsuperscript{2}CCLRC Rutherford Appleton Laboratory, UK
What is OPeNDAP?

- Enables seamless access to remote datasets
- Client/server model, with many different clients available
- URLs used to access datasets, with the data delivered via HTTP
- Data translation facility - original data may be stored in a wide variety of formats (netCDF, JGOFS, HDF...)
- Sub-sampling capabilities
DODS Index of /dods-ncep2/ gaussian_grid/air.2m.gauss

<table>
<thead>
<tr>
<th>Name</th>
<th>Last modified</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent Directory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1979.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1980.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1981.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1982.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1983.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1984.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1985.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1986.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1987.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1988.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1989.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1990.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1991.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1992.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1993.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1994.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1995.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
<tr>
<td>air.2m.gauss.1996.nc</td>
<td>05-May-2003</td>
<td>50M</td>
<td></td>
</tr>
</tbody>
</table>

Done
DODS Dataset Access Form

**Action:** Get ASCII | Binary Data Object | Show Help

**Data URL:**

**Global Attributes:**
- Conventions: "CF-1.0"
- title: "4x Daily NCEP/DOE Reanalysis 2"
- history: "created 2002/03 by Hoop (netCDF2.3)"
- comments: "Data is from NCEP/DOE AMIP-II Reanalysis (Remealys-2) (4x/day). Data interpolated from model (sigma) surfaces to gaussian"

**Variables:**
- **level:** Array of 32 bit Reals [level = 0.0]
  - level:
    - units: "m"
    - actual_range: 2., 2.
    - long_name: "level"
    - positive: "up"
    - axis: "z"
    - coordinate_defines: "point"

- **lat:** Array of 32 bit Reals [lat = 0.,93]
  - lat:
    - units: "degrees_north"
    - actual_range: 88.54199892, -88.54199982
    - long_name: "latitude"
    - standard_name: "latitude_north"
    - axis: "y"
The limitations of OPeNDAP

- Only allows sub-sampling - no other server-side manipulation of the data is possible
- Only allows references to single files
Extending OPeNDAP

- Use existing applications (Ferret, GrADS, IDL, Matlab, NCO...) to carry out server-side processing of the data
- Enables a wide range of operations to be performed on the data - arithmetic operations, averaging, calculating EOFs...
- Requires the OPeNDAP URL format to be expanded
- Requires only a trivial modification to existing OPeNDAP installations - just replace a few Perl modules
Expanding the OPeNDAP URL format

- Conventional OPeNDAP request:
  
  air.2m.gauss.1979.nc.dods?air[0:1:1459][0:1:93][0:1:191]

- Now calculate the average along the time axis:
  
  air.2m.gauss.1979.nc.dods?air[0:1:1459@ave][0:1:93][0:1:191]
Expanding the OPeNDAP URL format

- Calculate the average over multiple files:

  http://dods.acecrc.org.au/cgi-bin/nph-dods/_calc_average(inputs:air.2m.gauss.197?.nc)
  (params:input_var=air,output_var=average_air).dods?
  average_air[0:1:1459][0:1:93][0:1:191]

- Calculate the EOFs:

  http://dods.acecrc.org.au/cgi-bin/nph-dods/_calc_eofs(inputs:air.2m.gauss.197?.nc)
  (params:input_var=air,output_var=air_eof,tol=0.001).dods
Further developments

- Web processing service:
  - Asynchronous results delivery
  - Progress monitoring
  - Remote data input

- Web compute service:
  - Use JSDL (Job Submission Description Language) to specify resource requirements
  - Stage user data