

# Evidence of Implementation for DAP 2.0

James Gallagher\*

2004/08/09  
Revision 1.6

## 1 Implementations

The following list of implementations and operational sites is correct as of 8 August 2004. We know of other implementations and sites, but finding accurate information about them can be challenging since there is no requirement that users of the DAP communicate with us. Also, as is often the case with an open-source protocol, implementations sometimes come to light only after they have been developed and are used for some time.

The list below includes only those groups which have developed DAP software, as opposed to those groups that have developed software which uses the DAP. For the latter, see the list of 'Other groups that distribute DODS/OPeNDAP software'<sup>1</sup> at [www.opendap.org](http://www.opendap.org). Also note that information about servers is only included where the site has developed its own implementation of the DAP. For lists of known DAP-compliant servers, see 'OPeNDAP Datasets'<sup>2</sup> at the OPeNDAP web page. Section 2.1 discusses OPeNDAP's experience with its servers.

Because the DAP is intended to be language-neutral, it is important that there be implementations in different programming languages. The list below includes that information.

The Data Access Protocol (DAP) has both a client- and a server-side component, in the same sense that other protocols such as FTP or HTTP do. Because most people are interested in solely one or the other component, most of the implementations originating outside our group are for just the client or server part.

- The reference implementation for the DAP is maintained by OPeNDAP. It was originally written as part of the Distributed Oceanographic Data System (DODS) project and then became part of the National Virtual Ocean Data System (NVO DS) project. The implementation includes a library with a complete implementation of both the client- and server-components of the protocol as well as several client applications and a server capable of handling data stored/accessed in six different formats/APIs. Many government and university sites use this software; a partial list is available from the same location as the software. This software is written in C++ and is available at <http://www.opendap.org/>.

---

\*OPeNDAP, Inc., <[j.gallagher@opendap.org](mailto:j.gallagher@opendap.org)>

<sup>1</sup>[http://www.opendap.org/developers/third\\_party\\_software.html](http://www.opendap.org/developers/third_party_software.html)

<sup>2</sup><http://www.opendap.org/data/index.html>

- A second implementation of the DAP, written in Java, was started by personnel at JPL, working under subcontract as part of the DODS project. They built an implementation of the DAP for use in client development. Later, the implementation was extended to support servers as well, under subcontract with Oregon State University (OSU). Using the library, personnel at OSU developed a server for data held in relational (SQL) databases. Many government and university sites use this software; a partial list is available from the same location as the software. This software is written in Java and is available at <http://www.opendap.org/>. Note that this software is completely independent of the C++ reference implementation.
- One of the software ‘products’ we have built using the C++ DAP implementation is the ‘NetCDF Client Library.’ This is a library which uses the netCDF file access API to read data from DAP-compliant servers. This piece of software has been very successful and is used by many groups to make software they have already developed ‘network aware.’ In addition, Unidata, the developers of the original NetCDF software, have used the Java-DAP implementation to add a similar capability to the Java netCDF library software they distribute. This software has also been very successful and many Pure Java clients use it to access data from DAP-compliant servers. Some of the client applications which use one of these two Client Libraries are listed on the Software page<sup>3</sup> mentioned above.
- The GrADS Data Server (GDS) and OPeNDAP-enabled GrADS application are both distributed by the Center for Ocean-Land-Atmosphere Studies (COLA)<sup>4</sup>. The GDS is a DAP-compliant server which has been developed at COLA for internal use. However, they have decided to distribute the server under an open-source license. The OPeNDAP-enabled GrADS analysis program is also available from COLA. Both of these use software developed in-house at COLA as well as our software. Contact Brian Doty (< [doty@cola.iges.org](mailto:doty@cola.iges.org) >) for more information.
- The LDEO/IRI Climate Data Library<sup>5</sup> uses its own implementation of the DAP to read data from servers. The software is written in C. Contact Benno Blumenthal (< [benno@iri.columbia.edu](mailto:benno@iri.columbia.edu) >) for more information.
- A Python library for creating DODS/OPeNDAP<sup>6</sup> clients has been developed at the Universidade de São Paulo. Contact Roberto AF De Almeida (< [roberto@dealmeida.net](mailto:roberto@dealmeida.net) >) for more information.
- The High-Altitude Observatory (HAO) at UCAR has developed a server for CEDAR data. This software has been written in C++ and used our C++ DAP implementation as a starting point but differs in several significant ways. First, the server supports both HTTP and GridFTP access to data (the latter is part of the Earth System Grid II project). Secondly, because CEDAR data is stored and accessed in ways that differ considerably from those of the data types our server was built to handle, the HAO group has made significant extensions to our C++ class library. Many of those extensions have been, or will be, incorporated into the library we distribute, but they have their origin outside our group. Contact Jose Garcia (< [jgarcia@ucar.edu](mailto:jgarcia@ucar.edu) >) or Peter Fox (< [pfox@ucar.edu](mailto:pfox@ucar.edu) >) for information about this software.
- In the past various groups have talked to us about using the DAP as the backbone for an internal data system. It is very hard to track these developments because they mostly happen within the organizations. However, one such group which also provides access from outside their laboratory is at the Institut Pierre Simon Laplace (IPSL) in France. Contact ‘dodsipsl’ (< [dodsipsl@ipsl.jussieu.fr](mailto:dodsipsl@ipsl.jussieu.fr) >) with the subject “DODS-IPSL” for more information.

<sup>3</sup>[http://www.opendap.org/developers/third\\_party\\_software.html](http://www.opendap.org/developers/third_party_software.html)

<sup>4</sup><http://cola.iges.org/>

<sup>5</sup><http://ingrid.ldeo.columbia.edu/>

<sup>6</sup><http://opendap.oceanografia.org/>

An important point to keep in mind when looking at web references to the DAP is that in the past we did not make a strong distinction between the DODS and NVOADS projects and DAP. It is common to find references to the DAP which use the words 'DODS,' 'NVOADS,' or 'OPeNDAP' instead.

## 2 Operational Experience

Operational experience with the DAP can be divided into two categories: experience with software developed by OPeNDAP and experience with software developed by other groups.

### 2.1 OPeNDAP software

The OPeNDAP group has deployed data servers in use at more than forty locations which range from university sites (e.g., University of Rhode Island, University of Miami), government sites (e.g., NASA GSFC, NASA JPL, NOAA PMEL) to corporate sites (e.g., SAIC, UCAR).

The DAP is used to serve data stored at these sites and is used by clients to read data from those servers. The servers provide access to data stored in netCDF, HDF4, Matlab, and DSP files as well as data which can be accessed using either FreeForm or JGOFS (FreeForm is an API which can be used to describe ad hoc data file formats; JGOFS is a data system developed for the JGOFS project). Client applications include Ferret (using out netCDF client library), Matlab (using our command-line client tool) and the OPeNDAP Data Connector as well as significant set of other analysis programs.

The data made accessible using the OPeNDAP servers primarily fall under the oceanographic discipline, although not exclusively so. There are data sources which house satellite imagery, model output, and climatologies as well as point data from XBT, CDT and ships of opportunity. Estimates conservatively put the data volume at 10 TeraBytes.<sup>7</sup>

In some cases data are available using other protocols in addition to the DAP. In most cases where another access protocol is supported, it is FTP; files which are used by our servers are also easily made available for access using FTP.

The software implementing the DAP is maintained by OPeNDAP, a Rhode Island not-for-profit specifically created to main and extend the DAP and other related discipline-independent network data access technology.

For more information on OPeNDAP and the DODS/NVOADS servers, contact Peter Cornillon (< [pcornillon@opendap.org](mailto:pcornillon@opendap.org) >, 401.874.6283).

### 2.2 COLA/GDS/GrADS

The Center for Ocean-Land-Atmosphere Studies (COLA) provides access to climate and weather data using the DAP. Almost all of the data is gridded analysis and/or model output of atmospheric/ocean/land data, usually global. A small amount is in-situ or "station" data. About half is NCEP weather forecast model output, the other half is climate data and/or model output. Total data volume is estimated at in excess of 3 TeraBytes. COLA began serving data using the DAP in Sept of 2001. Access to data from the servers maintained by

---

<sup>7</sup>More realistic estimates are in the range of 20 to 30 TeraBytes, but but constantly changing data sources such as those as GSFC and JPL make this a very hard figure to pin down.

COLA is generally increasing; for the month of July 2004, 1500 unique IP address requested a total of 176 GigaBytes.

COLA has developed the GDS, a server built using both their own software and our software. The GDS adds extra capabilities in the form of server-side analysis. COLA provides this server freely to other sites for their own use.

COLA has combined their GrADS analysis program with our netCDF client library to create a client application which can read data from their servers (it can also read data for any DAP-compatible server).

All aspects of the DAP are used, with some extensions to the Constraint Expression syntax to facilitate the extra analysis capabilities of the GDS server.

COLA is committed to tracking changes in the DAP and maintain their own software. It should be pointed out that GrADS is a successful open-source analysis tool in its own right and the developers at COLA have a demonstrated track record of supporting that software for their user community.

For more information on COLA's use of the DAP, contact Brian Doty (< [doty@cola.iges.org](mailto:doty@cola.iges.org) >, 301.595.7000).

## 2.3 LAS

The NOAA Pacific Marine Research Laboratory (PMEL) has developed a web portal software package called Live Access Server (LAS) which uses the DAP to build and interconnect data-access web portals. Both data and visualizations are available from the sites. LAS makes available three to five TeraBytes of model, climate, oceanography, satellite, and real time tides-winds-wave data. PMEL is aware of 30+ institutions with publicly available LAS sites serving up data. Typical recent accesses are on the order of 400 Mbytes per month from the LAS operated at PMEL and estimated<sup>8</sup> at 2GB per month for the sum total of all LAS sites in operation.

Developers and researcher at PMEL have developed the LAS software and are in the process of building a server with analysis capabilities. In addition to LAS and the server under development, PMEL also distributes a DAP-enabled version of their Ferret analysis tool. Their involvement with the software using the DAP began in 1999.

All aspects of the DAP are used by the combination of DAP-enabled Ferret, LAS and the in-progress server.

PMEL is committed to tracking changes in the DAP.

For more information on PMEL's use of the DAP, contact Steve Hankin at PMEL (< [Steven.C.Hankin@noaa.gov](mailto:Steven.C.Hankin@noaa.gov) >, 206.526.6080).

## 2.4 LEDO/Ingrid

The Lamont-Doherty Earth Observatory (LDEO) at Columbia University hosts the IRI/LDEO Climate Data Library which contains over 300 datasets from a variety of earth science disciplines and climate-related topics.

The IRI/LDEO Climate Data Library (aka Ingrid) is implemented as a web portal which researchers use to browse and access data. Data are delivered both as files and using the DAP. Furthermore, Ingrid also reads data from many of the NVO DS servers, making it a client for those servers. The implementation is entirely in-house.

For more information on Ingrid, contact Benno Blumenthal (< [benno@iri.columbia.edu](mailto:benno@iri.columbia.edu) >).

---

<sup>8</sup>Jonathan Callahan, private communication, 8 Aug 2004.

## **2.5 Overlap in the Data Provided**

In the previous section, the data volumes contain some overlap. That is, some of the 3 TeraBytes of data accessible from COLA's GDS is also part of the NVOODS data source collection and some of Ingrid's data also appear as part of NVOODS and LAS. An effort to eliminate redundancy has been made, but it should be stressed that these numbers are estimates.