A National Oceanographic Partnership Program Award

HYCOM DATA ASSIMILATION AND WEB OUTREACH

Eric P. Chassignet
Meteorology & Physical Oceanography, University of Miami/RSMAS
4600 Rickenbacker Causeway, Miami, FL, 33149-1098
phone: (305) 361-4041 fax: (305) 361-4696 email: echassignet@rsmas.miami.edu

Grant Number: N000140210485
http://hycom.rsmas.miami.edu

Long-Term Goals

Implementation of an Information System that allows easy access to observational data and model output to (a) to members of the HYCOM consortium and for data assimilation code development, (b) the wider oceanographic and scientific communities, including climate and ecosystem researchers, and (c) the general public especially students in elementary and high schools.

Objectives

The main objective is to implement existing open source data distribution systems to make HYCOM Consortium generated data accessible from the Internet.

Approach

The basic task is to host an independent website that provides access to data generated by the HYCOM Consortium. This requires considerable amount of hardware and software resources to be tied together as a functional system. The hardware software and system architecture details are described below in detail.

Hardware

A PC-cluster running the Linux operating system is used to host the data, meta-data and the data distribution software. A picture of the system is shown in Figure 1. A PC-Cluster was chosen for the hardware architecture since it provides the best price to performance ratio for web application in contrast to conventional workgroup servers. The cluster consists of 13 dual processor (5 Intel PIII and 8 AMD Athlon 1900) nodes and a fileserver (CHI Corporation NAS 2000). The cluster is in a private network and one of the nodes is designated as the master and handles all communications with the outside world. It has direct attached storage capacity of ~ ½ Terabyte. This is in a RAID 5 configuration with individual enterprise class SCSI drives. Most of the data storage is provided by the fileserver, which has a capacity of approximately 2 Terabytes. The fileserver is a Network Attached Storage (NAS) appliance and is attached to the cluster through the Internet protocol. The arrangement has a total storage capacity of approximately 2.5 Terabytes and is easily scalable. Additional storage (NAS) can be added in a matter of minutes and without any downtime.

Software

The system is based on the conventional three-tiered distributed application architecture. The user typically interacts with a web interface that sends a request to the server. The application on the server retrieves the data from the fileserver and returns product back to the user. In implementing this system there are a number of choices in selecting the software. Two open source data management systems, namely OPeNDAP (http://unidata.ucar.edu/packages/dods/) or Open Project Network Data Access Protocols and LAS (http://ferret.wrc.noaa.gov/Ferret/LAS/) or Live Access Server, are the most commonly used applications in the ocean modeling community. OPeNDAP was developed jointly by MIT and URI and is a middleware that enables network access to data residing on OPeNDAP enabled servers to OPeNDAP enabled clients machines. OPeNDAP supports commonly used data formats and the middleware transparently translates data formats on transmission. LAS was developed at PMEL and provides a friendly user interface to geo-science data browsing, download and comparison over the Internet. The user can obtain the data in a variety of file formats including images and text files. Exhaustive technical details about these packages are available at the above sites. Both software packages are well supported and are available free of cost. These open source software packages were specifically developed for data access purposes and form the backbone of the system that is implemented at University of Miami. In addition to the above mentioned software packages numerous other helper applications and scripts were utilized to tie the system together. The various software packages used to build the system are listed below.

OPeNDAP (DODS) version 3.2.8

a) LAS version 5.2

b) Perl 5.6.0

c) NetCDF Libraries version 3.5.0

d) MySQL RDBMS version 1.3.3

e) Ferret verion 5.22

f) Apache version 2.1

g) Java version 1.4

h) Apache Tomcat Server

Several methods are currently employed by the ocean data community to logically aggregate individual snapshot flies in datasets consisting of timeseries of individual files. Two of these methods will be used in this installation to aggregate the data. The current implementation consists of using ferret descriptor files to aggregate datasets. This method limits the use of OPeNDAP clients to Ferret only and is not advisable in the long run. Therefore a second method of installing an OPeNDAP aggregation server will be used in conjunction with the Ferret descriptors files to provide aggregations services. Currently the development of the OPeNDAP aggregation server is in beta stages and is likely to be available as a production release by the end of 2002. This will be implemented as soon as it is available.

Work Completed (started in January 2002)

a) Hardware acquired and configured

b) Open source software customized and installed

c) Available model output converted to Network Common Data Format (NetCDF, http://unidata.ucar.edu/software) and installed on the server

d) Access enabled to two datasets (~150 GB).

Fig. 1: Picture of the PC-Cluster installed at the University of Miami

Results

OPeNDAP and Live Access Server software has been installed on a PC-cluster at the University of Miami to provide access to Ocean model datasets. This installation currently provides access to roughly 150 GB of data from the following data sets:

a) HYCOM output from a 1/3° North Atlantic Simulation

b) MICOM output from a 1/12° North Atlantic Simulation.

The data are available at http://hycom.rsmas.miami.edu/dataserver. Figure 2 shows a screen shot of the Live Access Server.

Future work will encompass the addition of datasets and participation in data sharing project for Global Ocean Data Assimilation Experiment (GODAE). The goal of the GODAE Data Sharing System is to deliver shared, distributed data to a variety of visualization and analysis tools that participating researchers are most familiar with.

One immediate use of the system will be the participation in a GODAE data sharing pilot project. The GODAE Data Sharing Pilot is a voluntary effort on the part of the GODAE participants to implement and evaluate components of a data sharing system amongst GODAE participants. The data sharing pilot will use the same software packages that were used to build the data distribution system.

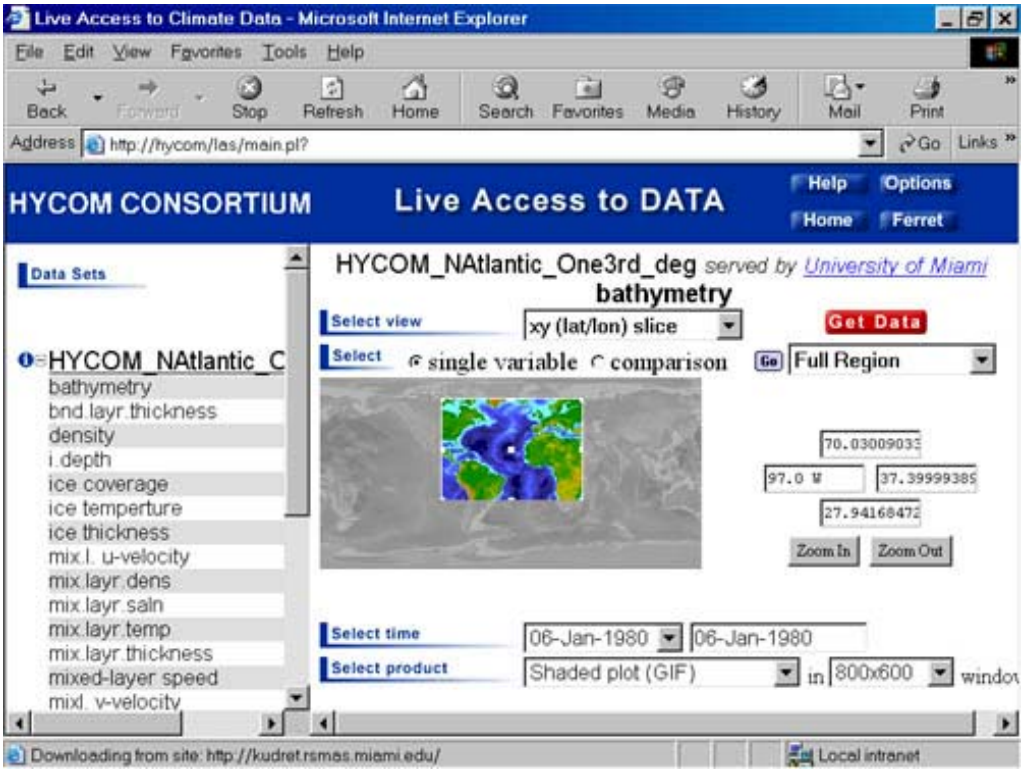

Fig 2. Screen shot of the operational Live Access Server (http://hycom.rsmas.miami.edu/las/).

A National Oceanographic Partnership Program Award

Impact and Applications

National Security

Generation of optimal estimates of the time-varying ocean state in support of the U.S. Navy's needs on synoptic time scales on the order of weeks to months and on spatial scales typically on the order of 10-1000 km (mesoscale).

What is the potential future impact on National Security or Homeland Defense?

Better knowledge of the oceanic circulation near the United States coastline.

Economic Development

Precise knowledge and prediction of ocean mesoscale features is used by the fishery and oil industry for an optimal use of their resources.

Science Education and Communication

Web-access to an up-to-date description of the world ocean currents.

Transitions

National Security

HYCOM is planned as the next generation eddy-resolving (~7 km mid-latitude resolution) operational global ocean nowcast/forecast system at the Naval Oceanographic Office (NAVOCEANO). Transition of the HYCOM-based global system from Research and Development to NAVOCEANO is planned for FY06.

Related Projects

This effort is part of a multi-institutional NOPP project which includes E. Chassignet (Coordinator), G. Halliwell, and A. Mariano (U. of Miami/RSMAS), T. Chin (JPL/U. of Miami), R. Bleck (LANL), H. Hurlburt, A. Wallcraft, P. Hogan, R. Rhodes, C. Barron, and G. Jacobs (NRL-Stennis), O.M. Smedstad (Planning Systems, Inc.), W.C. Thacker (NOAA/AOML), and R. Baraille (SHOM).