Addition of a Community THREDDS/RAMADDA Server System at Penn State
(Final Report for 2010 Unidata Equipment Awards)

Charles Pavloski and Arthur Person
The Pennsylvania State University

Abstract

The Pennsylvania State University (PSU), Department of Meteorology currently ingests, relays and processes most of the available Unidata Internet Data Distribution (IDD) data streams. Processing of the data is currently limited to raw and decoded gempak data sets for local use. Using funds from the 2010 Unidata Equipment Program, we have expanded our data distribution service through the addition of Unidata's Thematic Realtime Environmental Distributed Data Services (THREDDS), and the Repository for Archiving, Managing and Accessing Diverse DAta (RAMADDA) applications. We are now proud to announce the availability of this service for the purpose of distributing IDD and BUFKIT data to the greater Unidata community.

Our New Service to the Community

The PSU Meteorology Department has been a user of Unidata data since program inception and has been an IDD participant since 1997. We also serve the Unidata community through our Tier-1 level IDD relay service for CONDUIT data (hosted on idd.meteo.psu.edu).

This relay currently ingests and relays all major IDD data streams. We have taken the next logical step forward to augment our IDD service to the community with a THREDDS/RAMADDA data service. This new service includes current data and short-term archives (20 days) for a significant portion of the IDD data stream we currently ingest and is configured similarly to the Unidata IDD THREDDS servers.

This service also provides access to locally generated datasets, which currently consists of BUFKIT data files. We plan to add more locally-generated data streams to our server such as real-time, experimental forecast model output generated at Penn State and Pennsylvania climate data.

The new hardware is actually a pair of dual, quadcore, Xeon servers each with 16 GB of memory. One server has 11 TB of RAID 10 storage; the other server will be configured with another 11 TB of storage in the near future for a total of 22 TB. The file systems are cross-mounted on a private LAN to provide uniform data access from either server, and both servers are being configured to support http service using Apache Tomcat software. Currently, one server ingests/processes the IDD data stream, but in the future this will be split between both servers to even-out the work load.

Access Instructions

Available datasets include all the standard data available through the IDD as well as historical raw BUFR and decoded BUFKIT model sounding data for the NAM, GFS and RUC models from 2010 up to the present.

To examine these offerings in more detail, you may connect to the THREDDS server at http://tds.meteo.psu.edu:8080, or the RAMADDA repository at http://tds.meteo.psu.edu:8080/repository. If you encounter problems or have questions, please contact us at support@meteo.psu.edu.