5B.2

The LEAD testbed system at the Unidata program center: a medium term repository of meteorological data

Thomas Baltzer, Brian Kelly, Anne Wilson and Mohan Ramamurthy
Unidata Program Center, UCAR, Boulder, CO

1. INTRODUCTION

A key part of the Linked Environments for Atmospheric Dicsovery (LEAD) vision is that of providing seamless access to meteorological data for education and research purposes. That goal includes empowering the user community to assimilate, subset, download, visualize, mine and use a variety of meteorological observations, model output, radar data and satellite imagery through a simple, easy to use, and web-based LEAD portal (http://leadproject.org). Since LEAD is intended to support education and research in meteorology, it is necessary to facilitate case studies of significant meteorological phenomena by providing the ability to perform retrospective model runs and analysis of and comparison with the corresponding verification data.

In order to support these needs, the UCAR Unidata Program Center has built a relatively inexpensive medium term (~6 month, ~80 TB) data repository. This work has the additional benefit of providing a medium term repository of several data streams in the Unidata Internet Data Distribution system (http://www.unidata.ucar.edu/software/idd/) for the Unidata community. Another benefit to Unidata has been the opportunity to investigate scalability of certain technologies, including THREDDS (the Thematic Realtime Environmental Distributed Data Services – see:http://www.unidata.ucar.edu/projects/THREDDS/), and conduct stress tests on large volumes of data and longer term holdings.

This repository is available to both the LEAD and the Unidata communities for their purposes, but is not designed to be 100 percent guaranteed (see History below).

Corresponding author address: Thomas S.V. Baltzer PO Box 3000 Boulder Colorado, 80307: tbaltzer@ucar.edu

2. DESCRIPTION OF THE EFFORT

One of the primary Thrust groups in the LEAD effort is the Grid and Web Test Bed thrust. This group has the charge of establishing a set of Test Bed capabilities to support prototyping and development of overall LEAD system capabilities. Each of the institutions involved in this endeavor created test bed systems that focused on their own area of expertise. As a provisioner of meteorological data Unidata elected to develop a test bed system focused on an repository of these data along with providing on-line access to that repository. Additionally, we included a small processing cluster to work with tools that the LEAD project incorporates.

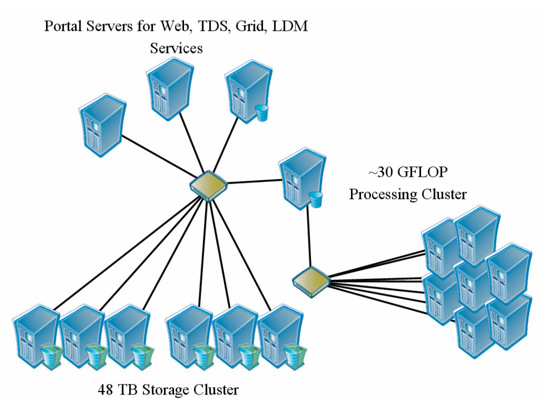

Figure 1: Testbed Configuration at Unidata/UCAR

Figure 1 shows the overall hardware Infrastructure put together in support of the Unidata LEAD Test Bed System. A significant desire was to create an inexpensive yet functional test bed system. As it happened, desktop systems around Unidata were being replaced at the time the test bed system was being developed, so we were able to repurpose a lot of hardware to this effort. Ultimately the system was created as follows:

-Less than $70 K in newly purchased hardware
-Approximately $10K of repurposed hardware
-Approximately 500 person hours of design and building time

Research Objectives

The primary research objectives of this work are to 1) Provide LEAD with a medium term repository of meteorological data for use in near real time analysis and processing as well as retrospective analysis and processing, 2) Provide an HPC multi-processor environment for the LEAD team at Unidata to experiment with LEAD applications (e.g. WRF), 3) Provide the THREDDS group with an example case of using the technology for a medium term (~6 month) repository, 4) Provide the Unidata Community with access to a Medium Term Repository of Internet Data Distribution (IDD) data.

Medium Term Data Repository

Unidata's motherlode system is a well known short term (7 day) data holding used extensively by the community. It provides cataloging of the data via the THREDDS Data Server (TDS) and access to the data via several protocols (ADDE, OPeNDAP, HTTP, etc.)

The LEAD project chose to use the TDS with OPeNDAP and HTTP access protocols and add in a GridFTP access to the stored data. Figure 2 shows how the Unidata LEAD testbed system is configured.

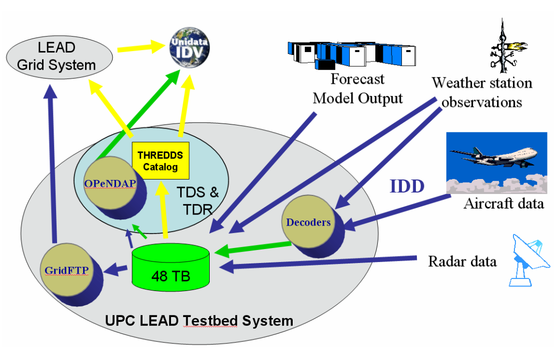

Figure 2: Data Configuration at the UPC LEAD Testbed

Since the repository is intended to serve two audiences (LEAD and the Unidata Community) it was necessary to establish two different THREDDS cataloging hierarchies, one for each community. Thus the LEAD community is supported by the catalog found at:

http://lead.unidata.ucar.edu/thredds/topcatalog.html

and the Unidata Community is supported by the catalog found at:

http://lead.unidata.ucar.edu/thredds/catalog.html

Each of these catalogs has an xml equivalent which the Unidata Integrated Data Viewer (IDV) reads natively. This works out nicely for the LEAD project in that the IDV was selected as the visualization tool for LEAD.

Two different catalogs are needed since LEAD uses/supports only a subset of the IDD datasets available and additionally has datasets that are unique to LEAD – for example, the Oklahoma University LEAD partner is running a 10km ADAS assimilation hourly for the LEAD project.

Dataset	# Days
NAM 40 km CONUS	300
ADAS 10 km CONUS	45
NEXRAD Level II	183
Surface Observations	600
Profiler Data	600
Upper Air Observations	130
UPC Steered Regional WRF	130
Profiler	600
All NAM datasets	300
DGEX	220
RUC	220
NEXRAD Level III	200
HRS	160
Satellite	160
ECMWF	130
CMC	130
UKMET	130
GFS 0.5 Deg	120
RUC2	105

Table 1: Current Dataset Holdings in the Repository

Table 1 shows the current dataset holdings found on the repository. The number of days represents from the present, back in time. The table is divided into two parts to inform the reader about the datasets available exclusively through the LEAD catalog hierarchy. Thus far, LEAD is not leveraging all datasets found on the repository, but only those found above the blank row in the table.

Additionally, the repository is home to a number of case studies including Katrina and the Greensburg F5 tornado. Development of the THREDDS Data Repository (TDR) leverages repository as does by extension the Next Generation Case Study project is also making use of the repository for the purpose of making more such case studies.

At present the storage repository is composed of approximately 24 TB of data and served approximately 4.2 TB of data to the communities it serves last year. We continue to work to expand upon those numbers.

The LEAD project leverages this repository by regularly cross-walking the metadata found in the THREDDS catalog designed for LEAD's use into the LEAD Metadata Schema (LMS) and placing that in the LEAD database.

The Storage Cluster

In order to create the inexpensive storage cluster in such a way as to assure some degree of reliability a number of hardware and software technologies were brought together. The cluster consists of 6 systems each of which (figure 3) is composed of twenty-four (24) 400 GB drives, divided into two eleven column RAID 5 arrays and two hot spares in order to form two 4TB LUNs (Using bcraid) which are striped together using a RAID hardware card.

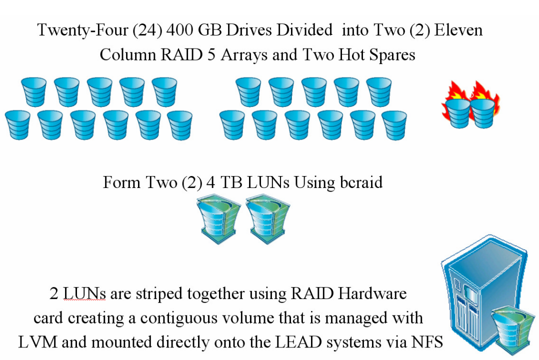

Figure 3: Composition of the Storage Nodes

This results in an 8TB contiguous volume that is cut up with a Logical Volume manager and NFS mounted directly onto the test bed systems in need of access (LDM, FTP and gridFTP to load, TDS and gridFTP to serve).

Each of the storage nodes consists of the following hardware:

- One (1) Guanghsing GHI-583 5U Case with 24 hot swapable SATA trays and a 1000W 2+2 power supply
- One (1) Tyan Thunder K8SD Pro Motherboard with Dual Opteron CPUs, Four 64-bit 133/100 Mhz PCI-X Slots andTwo Gigabit Ethernet ports
- One (1) AMD Opteron 242 Processor 1.6 Ghz CPU
- Three (3) Broadcom RAIDCore BC4853 with Eight SATA ports, Controller spanning and Advanced raid
- Twenty-Four (24) Seagate Barracuda ST3400832AS 7200 RPM 400GB SATA Drives

3. History

Assuring Reliability

When the repository was initially set up, the vision was to mirror all the data to doubly assure that no data would be lost. However, given that the system is receiving over 150,000 data products totaling over 4GB per hour, we found that the mirroring function overwhelmed the Input/Output channels on the storage nodes. This resulted in the loss of data products before they were even stored in the repository. As such, this approach was abandoned.

The belief was that the RAID configuration of the storage nodes would assure that no data would be lost. The RAID technology tolerates the loss of a hard disk in the system by having "hot spare" disks that can be used to replace a disk that goes bad. However, when a disk does go bad, the RAID must be "rebuilt" using the hot spare disk. During this time, performance of the system is compromised, as is the fault tolerance. Should a second disk in the system go bad, all data that has not been rebuilt, will be lost. As fate would have it, this situation occurred on the testbed.

In this situation, the team was able to recover some of the data and put it onto another of the storage nodes, but without guarantees of reliability. The worst part of this was that data the LEAD project needed to assure users' ability to perform retrospective weather prediction runs was lost. As such the team is working on saving those critical datasets on two different storage nodes so that should the fickle finger of fate strike again, users will still be able to perform retrospective runs.

As the repository was fist being filled, the team noticed a degradation of performance on the storage node responsible for holding the Level III radar data.

Future Directions

Currently the repository is at just over 50% of capacity. The expansion of datasets has been approached slowly in order to observe the impacts on reliability and speed of delivery. It is intended to maintain this approach to filling the archive to at least 75% of capacity. So long as reliability and speed of delivery be maintained, the expansion will continue.

This will provide further experience with the use of the Unidata TDS and TDR technologies in handling larger holdings of data.

New datasets will be added to the repository as well. For example, the LEAD project is ready to begin working with 12km tiled NAM datasets for use as initial and boundary conditions in workflows for mesoscale modeling.

4. ACKNOWLEDGEMENTS

LEAD is a Large Information Technology Research (ITR) Grant funded by the National Science Foundation under the following Cooperative Agreements: ATM-0331594 (University of Oklahoma), ATM-0331591 (Colorado State University), ATM-0331574 (Millersville University), ATM-0331480 (Indiana University), ATM-0331579 (University of Alabama in Huntsville), ATM03-31586 (Howard University), ATM-0331587 (University Corporation for Atmospheric Research), and ATM-0331578 (University of Illinois at Urbana-Champaign, with a sub-contract to the University of North Carolina).

5. REFERENCES

The Linked Environments for Atmospheric Discovery (LEAD) http://leadproject.org

THREDDS Catalog Specification http://www.unidata.ucar.edu/projects/THREDDS/tech/catalog/InvCatalogSpec.html

OPeNDAP Project http://www.opendap.org/

The Integrated Data Viewer (IDV) http://www.unidata.ucar.edu/software/idv/

The Weather Research & Forecasting (WRF) Model http://www.wrf-model.org/index.php