Use of TDS to support the LEAD Project

Tom Baltzer

For Unidata Workshop Summer 2007

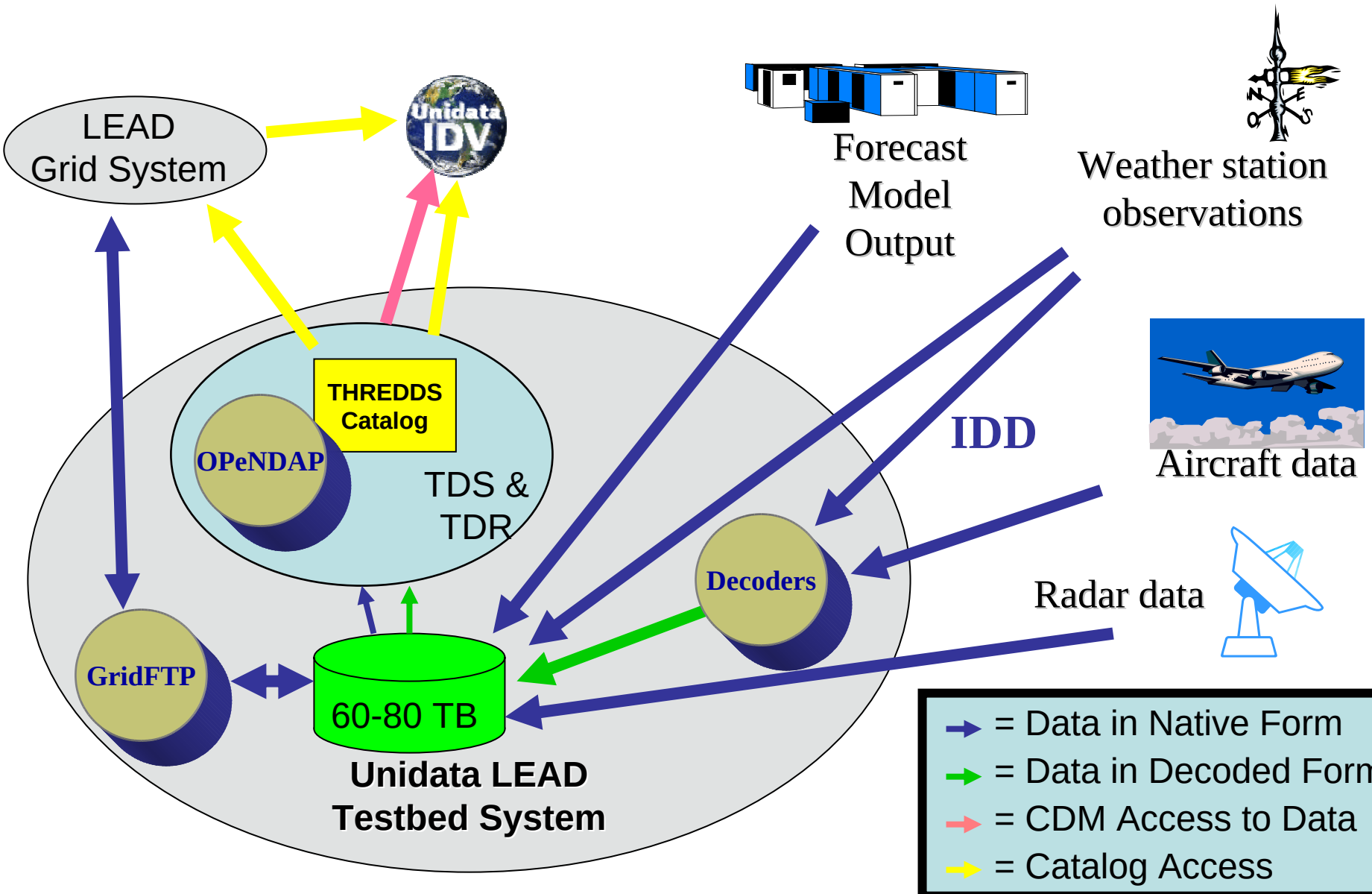
Linked Environments for Atmospheric Discovery (LEAD)

- Produce a web service and portal based, scalable framework for handling meteorological data and model output:
  - Identifying, accessing, preparing, assimilating, predicting, managing, analyzing, mining, visualizing
  - Independent of data format and physical location
- Dynamically adaptive workflows for model runs and steering of sensors
- Funded by NSF Large Information Technology Research (ITR) award

LEAD Portal


LEAD Portal - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://portal.leadproject.org/gridsphere/gridsphere;jsessionid=7E8A94E7EAE5F691D852E217920AD9EA?cid=login
Customize Links Free Hotmail Windows Marketplace Windows Media
LEADPORTAL
LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY
SPONSORED BY THE NATIONAL SCIENCE FOUNDATION
HOME MY WORKSPACE ABOUT LEAD DATA SEARCH EXPERIMENT VISUALIZE EDUCATION RESOURCES HELP
Welcome
WELCOME TO THE LEAD PORTAL
Linked Environments for Atmospheric Discovery (LEAD) makes meteorological data, forecast models, and analysis and visualization tools available to anyone who wants to interactively explore the weather as it evolves. The LEAD Portal brings together all the necessary resources at one convenient access point ... read more
WELCOME Tom Baltzer
LOGOUT
MyLEAD My Profile
QUICK LINKS
Live Weather
LEAD Grid
Glossary
Website Help
Frequently Asked Questions
FEATURES FOR ANYONE INTERESTED IN THE WEATHER
Researchers With university, government, or industry affiliations GET FEATURES
Educators At college and university level, high school, or middle schools GET FEATURES
Students At graduate, undergraduate, middle and high school levels GET FEATURES
Visitors Newcomers and the curious GET FEATURES
THE LEAD TEAM
Colorado State University
HOWARD UNIVERSITY
INDIANA UNIVERSITY
MILLERSVILLE
UAH
UCAR
ILLINOIS
OKLAHOMA
UNC
POPULAR TOOLS
Visualize Weather Data
Integrated Data Viewer | MORE
Make a Forecast or Analysis
Experiment Builder | MORE
Access Weather Data
Geographic Region Search | MORE
Done
portal.leadproject.org

The UPC LEAD Test Bed



LEAD Test Bed is supporting LEAD and Unidata Communities

- Need to support two different top level catalogs
  - For LEAD project
    - http://lead.unidata.ucar.edu:8080/thredds/topcatalog.html
    - http://lead.unidata.ucar.edu:8080/thredds/topcatalog.xml
  - For Unidata community
    - http://lead.unidata.ucar.edu:8080/thredds/catalog.html
    - http://lead.unidata.ucar.edu:8080/thredds/catalog.xml

    Shooting for 6 month archive of IDD data
    - Scalability testing for TDS

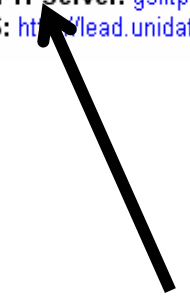
GridFTP Service

- Follow hierarchy from topcatalog
  - LEAD Testbed Catalogs ->
  - Unidata LEAD Testbed ->
  - LEAD IDV Viewable Catalogs ->
  - NAM Model Grids ->
  - CONUS 40 km (conduit)
- And select a given file

GridFTP “Service”

Access:

1. **OPENDAP:** http://lead.unidata.ucar.edu:8080/thredds/dodsC/LEAD/model/NCEP/NAM/CONUS_40km/conduit/NAM_CONUS_40km_conduit_20070718_1800.grib1
2. **HTTPServer:** http://lead.unidata.ucar.edu:8080/thredds/fileServer/LEAD/model/NCEP/NAM/CONUS_40km/conduit/NAM_CONUS_40km_conduit_20070718_1800.grib1
3. **GridFTPServer:** gsiftp://lead2.unidata.ucar.edu/gridftp/LEAD/model/NCEP/NAM/CONUS_40km/conduit/NAM_CONUS_40km_conduit_20070718_1800.grib1
4. **WCS:** ht…/lead.unidata.ucar.edu:8080/thredds/wcs/LEAD/model/NCEP/NAM/CONUS_40km/conduit/NAM_CONUS_40km_conduit_20070718_1800.grib1



GridFTPServer is a special Access provided for LEAD

GridFTP Service

```
[tomcat@lead4 lead]$ more lead_nam_model.xml
<?xml version="1.0" encoding="UTF-8"?>
<catalog xmlns="http://www.unidata.ucar.edu/namespaces/thredds/InvCatalog/v1.0" xmlns:xlink="http://www.w3.org/1999/xlink" name="NCE
P Model Data" version="1.0.1">
  <service name="latest" serviceType="Resolver" base="" />
  <service name="all" serviceType="Compound" base="">
    <service name="ncdods" serviceType="OPENDAP" base="/thredds/dodsC/" />
    <service name="HTTPServer" serviceType="HTTPServer" base="/thredds/fileServer/" />
    <service name="GridFTP" serviceType="GridFTPServer" base="gsiftp://lead2.unidata.ucar.edu/gridftp/" />
    <service name="wcs" serviceType="WCS" base="/thredds/wcs/" />
  </service>
  <dataset name="North American Model">
    <metadata inherited="true">
```

Setup new service type and GridFTP server shares disks with TDS server
Key – GridFTP works with URLs

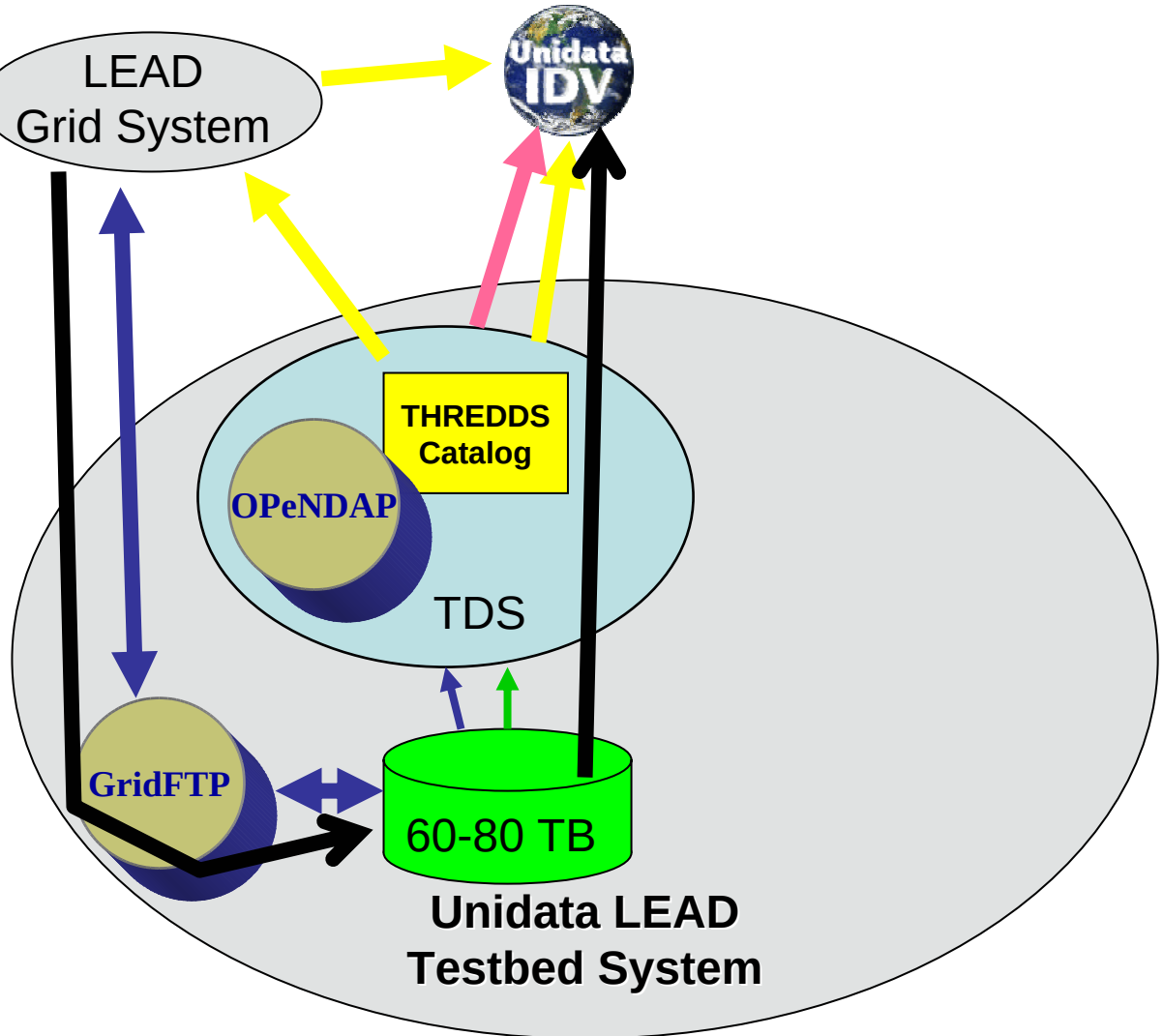
```
    <datasetScan name="CONUS 40 km (conduit)" collectionType="ForecastModelRuns" harvest="true" ID="NCEP/NAM/CONUS_40km/conduit"
path="LEAD/model/NCEP/NAM/CONUS_40km/conduit" dirLocation="/data/pub/native/grid/NCEP/NAM/CONUS_40km/conduit/" filter=".*\.grib1$
" addDatasetSize="true" addLatest="true">
```

Key - paths need to match!

```
[root@lead2 conduit]# pwd
/gridftp/LEAD/model/NCEP/NAM/CONUS_40km/conduit
[root@lead2 conduit]#
```

LEAD Results

LEAD Systems (including TeraGrid) will generate result files that need to be stored, cataloged and made accessible



LEAD Result Files

- Workflow system will deposit files on the Unidata LEAD testbed via GridFTP
- These files are considered “Private” – that is, they belong to a LEAD user who does not (yet) wish to share them
  - myLEAD (private catalog) is only place where URL is to be registered

LEAD Result Files

- How do we serve via TDS but keep private?
- Security through obscurity
- Catalog is put into threddsConfig.xml

```
<catalogRoot>lead/workshop_wrf_model.xml
</catalogRoot>
```

Demonstration of LEAD Use Case

Rainstorms over OK on 7/9/2007

Questions?