Unidata: An Overview

Tom Yoksas

Unidata Program Center / UCAR
Boulder, Colorado USA
http://www.unidata.ucar.edu
Unidata - Principle Activities

- **Providing Tools** to visualize, analyze, organize, receive, & share data
- **Facilitating Data Access** to a broad spectrum of observations & forecasts (most in real-time)
- **Supporting Faculty** who use Unidata systems at universities
- **Building a Community** where data, tools, & best practices in education/research are shared
Unidata Community

Unidata user community is interdisciplinary - 2/3rd of sites have users outside atmospheric sciences

We are moving from an era of data provision towards one in which data- and related web-services are important;

Multidisciplinary integration and synthesis are emphasized.
Tools and Support Are Central

- Enhance and distribute software developed by others
  - Meteorological display and analysis tools from UW-Madison (McIDAS-X), National Weather Service/NCEP (GEMPAK), etc.
  - Remote access technologies: OPeNDAP (U of RI, NASA, and others), ADDE (UW-Madison)
- Develop software in-house
  - Widely used tools for managing scientific data
    (e.g., LDM, netCDF, UDUNITS, data decoders, etc.)
  - Java-based tools (IDV Framework built on top of VisAD) for 2D, 3D, & 4D visualization and next-generation collaborative data analyses
- Build systems from the software we support
  - Internet Data Distribution (IDD) system
  - THematic Realtime Environmental Data Distributed Services (THREDDS)
- Support the use of software by offering training, consultation, bug fixes, and upgrades
Unidata Visualization Software

GEMPAK

McIDAS-X

IDV
IDV Visualization Examples

Sea-level Pressure and Upper-level Jet

Upper-mantle convection

NO₂ concentration

S-POL Radar Thunderstorm Cross-section
Internet Data Distribution System

- Initiated in the mid-1990s in response to weather-data ingest challenges:
  - Solar occultation data loss
  - Terrestrial interference
  - Campus *beautification* committees

- Event-driven network of cooperating Unidata Local Data Manager (LDM) servers interconnected by TCP/IP Ethernet

- Built to realize a communications goal laid out in the earliest Unidata planning documents (Cooper, 1985)
  - Active use of local-area and national network infrastructure
  - Allow for multi-way sharing of data including locally-held datasets

- Evolved in lock-step with national and international networking capabilities
Internet Data Distribution System Concept

Sharing data from multiple sources using cooperating LDMs
SuomiNet GPS Met and Geodetic Sites

Unidata’s LDM is used for real-time data transport
Real-time Data Flows

In the Beginning...  

“a dizzying volume of information – on the order of 100 MB/day, aggregate”  
(Davis and Rew, 1990)

<table>
<thead>
<tr>
<th>Service</th>
<th>Data Rate (Bits Per Second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAFAX</td>
<td>Analog</td>
</tr>
<tr>
<td>DIFAX</td>
<td>2400</td>
</tr>
<tr>
<td>Watches &amp; Warnings</td>
<td>1200</td>
</tr>
<tr>
<td>U.S. Surface/Upper-Air</td>
<td>4800</td>
</tr>
<tr>
<td>NMC &amp; ECMWF Grids</td>
<td>4800</td>
</tr>
<tr>
<td>Wisconsin Channel</td>
<td>9600</td>
</tr>
<tr>
<td>International (GTS) Data</td>
<td>1800</td>
</tr>
<tr>
<td>FAA 604</td>
<td>1200</td>
</tr>
<tr>
<td>Lightning Data &amp; Others at Non-Discounted Prices</td>
<td></td>
</tr>
</tbody>
</table>

LDM-6 Internet2 bandwidth use
Typically 21 TB/week & growing
Continue to push LDM-6 to its technological limit

Unidata top-level IDD relay cluster

- Built using IP Virtual Server built in current Linux distributions
- Mixed OS environment
  - 2 distributed directors (64-bit FC6 Linux)
  - 4 distributed real data servers (64-bit Linux Web 100 kernels)
  - 2 data collectors (Linux and FreeBSD)
- Demonstrated relay of sustained >500 Mbps (5.4 TB/day) during three day “live” trial; peak rates exceeded 900 Mbps
- Ability to relay data limited only by underlying network bandwidth (1 Gbps)
- Services > 400 concurrent downstream connections with zero added latency

Cluster-based relay installations:

- NOAA/GDS (formerly FSL)
- Penn State University (in progress)
- TAMU (?)
Real-Time Data Distribution

Unidata’s reach is global. IDD use:
Over 200 sites on six continents and growing!

http://www.unidata.ucar.edu/software/idd/rtstats
**LDM Technology Transfer**

**US NWS**
- CRAFT: NEXRAD Level II
- CONDUIT: NCEP High resolution model output

**Korean Met. Admin.**
- distribute all data to forecast offices

**Spain**
- deliver MSG data to offices

**Brazil**
- IDD-Brasil
The Impending Data Deluge

More Data and New Data Sources
  - Both GOES-R and NPOES will have data rates 30-60 times the current
- METOP (Europe polar orbiter(s))
- Raw data rate: 3 terabytes per day
- Global, coupled models at a grid spacing of 1-5 km, integrated for multi-decades
- NCAR Global WRF model for use in Weather and Climate research
- NEXRAD Level II expansion by 27x
- TIGGE
- SCOOP
- AMPS
- New initiatives…
Historic Data Paradigm

Unidata user running GEMPAK, IDV, McIDAS

Local data decoded into application specific formats

Application specific protocols

Forecast
Model Output

Weather station observations

Satellite imagery

Radar data

Lightning, aircraft, GPSmet, etc.
Thematic Real-time Environmental Distributed Data Servers (THREDDS)

- Combines IDD “push” with several forms of “pull” and DL discovery
- About 25 data providers are partners in THREDDS

To make it possible to publish, locate, analyze, visualize, and integrate a variety of environmental data

Connecting People with Documents and Data

User applications: e.g., McIDAS, IDV, LAS, IDL, MatLab...

OPeNDAP, ADDE, THREDDS DS

Hydrology Data, e.g.

Geophysical Data, e.g.

Satellite Imagery...

IDD

IDD

IDD

NSDL
Digital Library for Earth-System Education

DL interchange protocol

THREDDS Middleware

Discovery and Publication Services

Data Catalog Services

Discovery and Publication Tools

Analysis and Visualization Tools

Data Services

Discovery and Publication Tools

Analysis and Visualization Tools

Data Services

Documents

Data

People

Hydrology Data, e.g.

Geophysical Data, e.g.

Satellite Imagery...

OpenDAP, ADDE, THREDDS DS

NSDL
Digital Library for Earth-System Education

DL interchange protocol

THREDDS Middleware

Discovery and Publication Services

Data Catalog Services

Discovery and Publication Tools

Analysis and Visualization Tools

Data Services

Documents

Data

People

Hydrology Data, e.g.

Geophysical Data, e.g.

Satellite Imagery...

OpenDAP, ADDE, THREDDS DS

NSDL
Digital Library for Earth-System Education

DL interchange protocol

THREDDS Middleware

Discovery and Publication Services

Data Catalog Services

Discovery and Publication Tools

Analysis and Visualization Tools

Data Services

Documents

Data

People
NetCDF / HDF5 Merger
LEAD: A Large Grid Computing Project

Linked Environments for Atmospheric Discovery

- Identify, Access, Assimilate, Predict, Manage, Mine, and Visualize a broad array of meteorological data and model output, independent of format and physical location
- Develop a range of Grid and Web Services for dynamic, on-demand, end-to-end weather prediction
- Institutions:
  - U. Oklahoma
  - Unidata
  - U. Alabama
  - U. Illinoi
  - U. Indiana
  - Millersville U
  - Howard U.
  - Colorado State U.
LEAD Prototype 4

Employ components of WRF prediction as a series of linked web services in a Grid Environment.
Contact Information

Unidata HomePage:
http://www.unidata.ucar.edu

Unidata LDM HomePage
http://www.unidata.ucar.edu/software/lmd

Unidata IDV HomePage
http://www.unidata.ucar.edu/software/idv

Unidata User Support:
support@unidata.ucar.edu