Director's Report

Unidata Users Committee Meeting

31 March 2014

Boulder, CO

Mohan Ramamurthy
Unidata Program Center
UCAR Community Programs
Boulder, CO

Unidata 5-Year Proposal

- The proposal was well reviewed and the NSF Review Panel unanimously recommended it for funding.

  - "The rational for Unidata's existence has not diminished with time. Rather, as noted in the proposal, the atmospheric science and geosciences communities need the services and innovations provided by Unidata as much now as when Unidata was first established."

- The proposed vision to transition Unidata's data services to a cloud environment was strongly endorsed.

  - "Cloud-based approaches enable a wider range of users to fully participate and make use of Unidata offerings by decreasing the infrastructure needs on the end of the user. This will even enable high schools and community colleges, without dedicated Linux labs, to make use of the data analysis and visualization resources for teaching and research."

- Unidata should develop and implement online software training ASAP.

Unidata 2018 Award Delays

- Since the total award amount exceeded $10M, it had to go through an automatic pre-award audit by NSF's Cost Analysis and Audit Resolution branch "to verify that costs are reasonable, allowable, and allocable in accordance with federal regulations and NSF guidance."

- The CAAR audit has resulted in funding delays, and the UPC is now operating under a pre-spending authorization.

- We hope to get all of the issues resolved soon so that NSF can go ahead and make the award.

Security Breach

- On Feb. 20, there was a security breach of the machine hosting the Unidata web server.

- A hacker had found and exploited an old CGI script on our web site; this script is not in wide-enough use to show-up in industry-standard vulnerability reports, but it is common enough to be used in the OPeNDAP community.

- The hacker apparently found the script on the Unidata web site through a Google search.

- The hacker was, however, able to download some backup files of our web site's backup database containing 47,333 registration entries for Unidata users, including encrypted user passwords.

- UPC staff were alerted while the exploit was in progress and the errant script was removed shortly thereafter.

- Users were immediately notified and urged to change their passwords on accounts at other locations if they had used the same password. All user passwords were reset on the Unidata web server.

New Staff

- Christian Ward-Garrison joined UPC's THREDDS team in December.

- Ryan May joined the UPC software development staff in January.

- Josh Young joined as our Community Services Manager on March 10th.

Other Staff Transitions

- Russ Rew began his two-year phased retirement last October; He is working 50%.

- Ben Domenico has reduced his time to 25% in January.

- Linda Miller retired on 30 November. **This is the first governing committee meeting Linda without Linda's presence in 23 years!**

Real-time Data Flows

Unidata IDD Topology 111025/1350
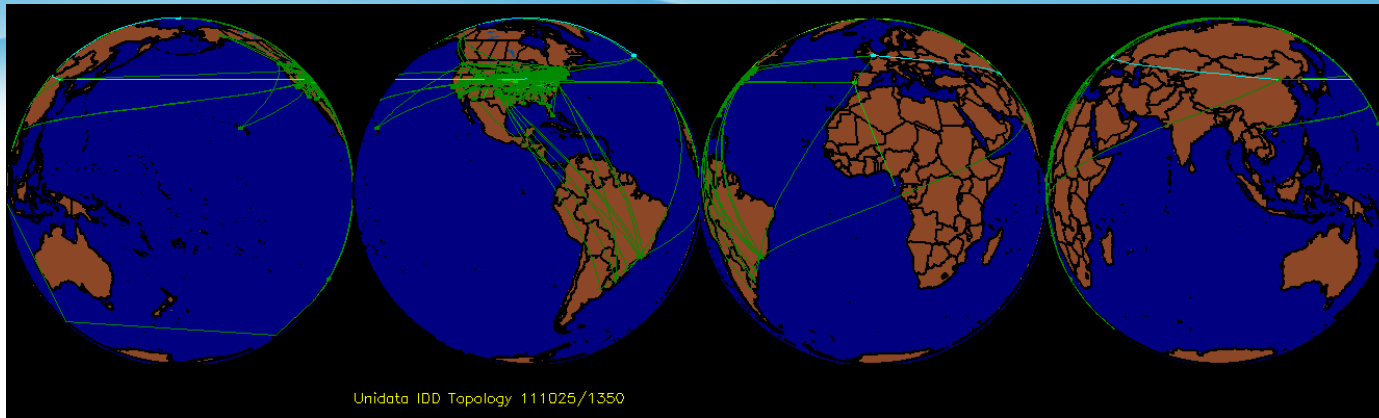
- ❖ About 530 machines at ~230 sites are running LDM-6 and reporting real time statistics. These numbers have not changed much in the past 4-5 years, but the volume data ingested keeps growing. Many organizations are using the LDM but not reporting stats.
- ❖ UPC's IDD Cluster relays data to about 700 downstream connections. Average data output: ~13 TB/day or 1.1 Gbps! Peak rate exceeds 2.2 Gbps!
- ❖ Data input to the cluster is ~15 GB/hr because of WSR 88-D Dual-polarization upgrades and the addition of more model output to CONDUIT.
- ❖ WSR 88-D Level II and CONDUIT remain the top two data streams based on volume.

Pilot Cloud Efforts

- We are taking initial steps toward creating a "Motherlode-lite" type data services on AWS;

- We are now moving a few data streams to the cloud, generating products (radar and satellite) there, and streaming the resulting products back;

- We would like to gradually expand the service offerings and entrain a few users to do some testing;

- Costs for moving data out of the cloud is a major consideration in how this will be set up, but prices are dropping rapidly.

Other Cloud Efforts

- Unidata has been awarded Windows Azure cloud resources, as part of the Microsoft "Windows Azure for Research Award Program";

- The award includes 32 compute instances, 10TB of storage, and 2 TBs network egress/month;

- The award was made based on a proposal to bring the Integrated Data Viewer (IDV) to modern devices via *Application Streaming*.

Application Streaming

- **Application Streaming** allows for a legacy program to run remotely, as if it were running natively on a client device.

- Compared to remote desktop software, only the application is streamed.

- Little-to-no re-engineering of legacy software is required.

- Because this implementation will be cloud based, dedicated IDV instances can be provisioned on an as-needed basis.
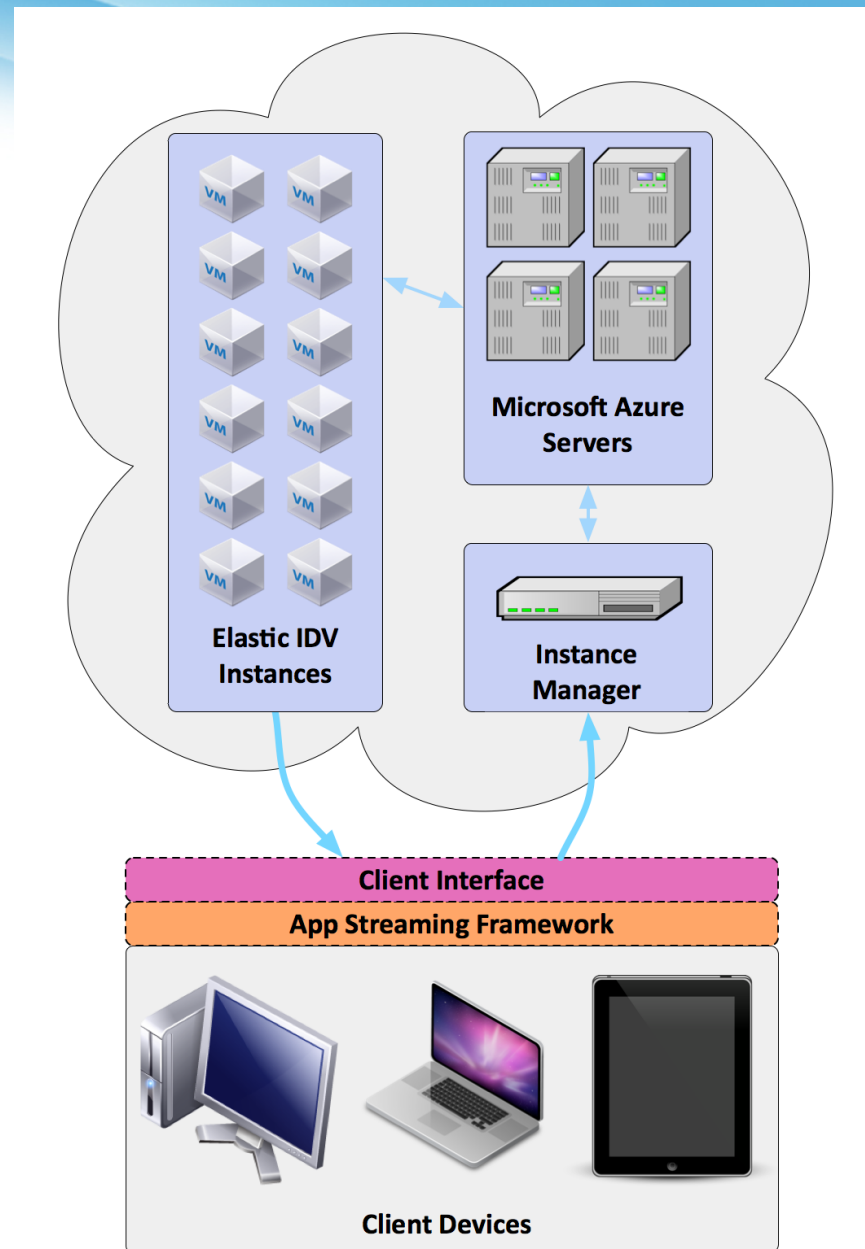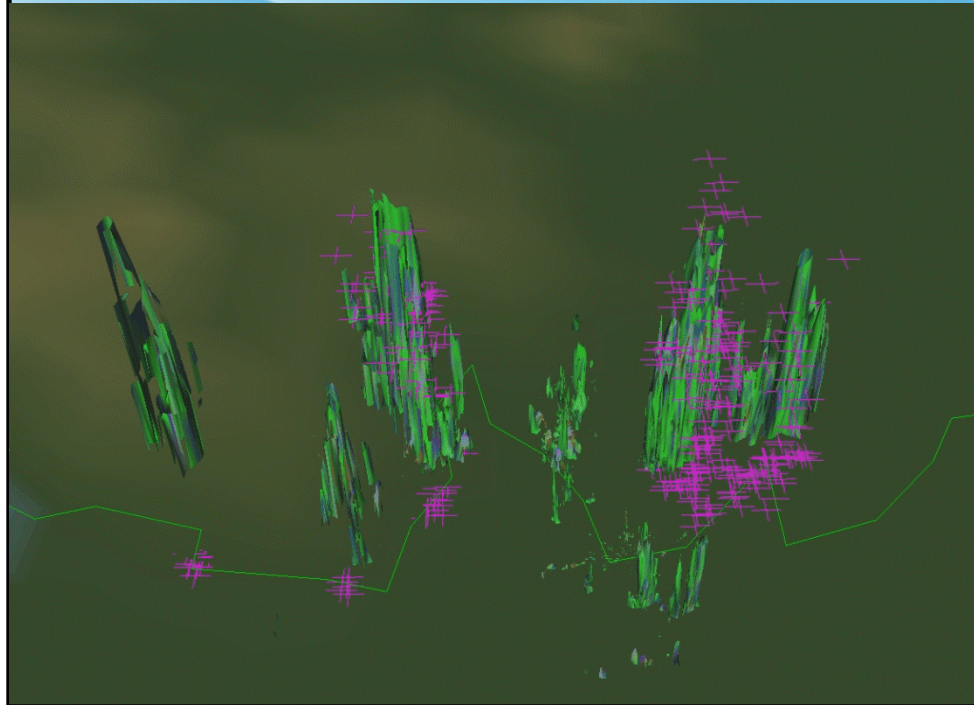
Servers at NCEP Project

- An initiative to result from the Open Weather and Climate Services concept/white paper
- Endorsed by NOAA, NWS, and NCEP
- Bringing processing to model-generated data
  - S erver-side processing is needed to effectively use high-volume model output
- A time-limited experimental data processing center co-located at the new NCEP supercomputing facility in Reston, VA
  - It will connect directly with NCEP supercomputer file system via high-speed fiber cable, enabling mirroring of more complete model data sets
  - Computers installed by participating community members will be able to connect directly to the mirrored data file system, allowing value-added processing to occur in the new facility and thereby lowering the volume of data moved to participants' home organizations.
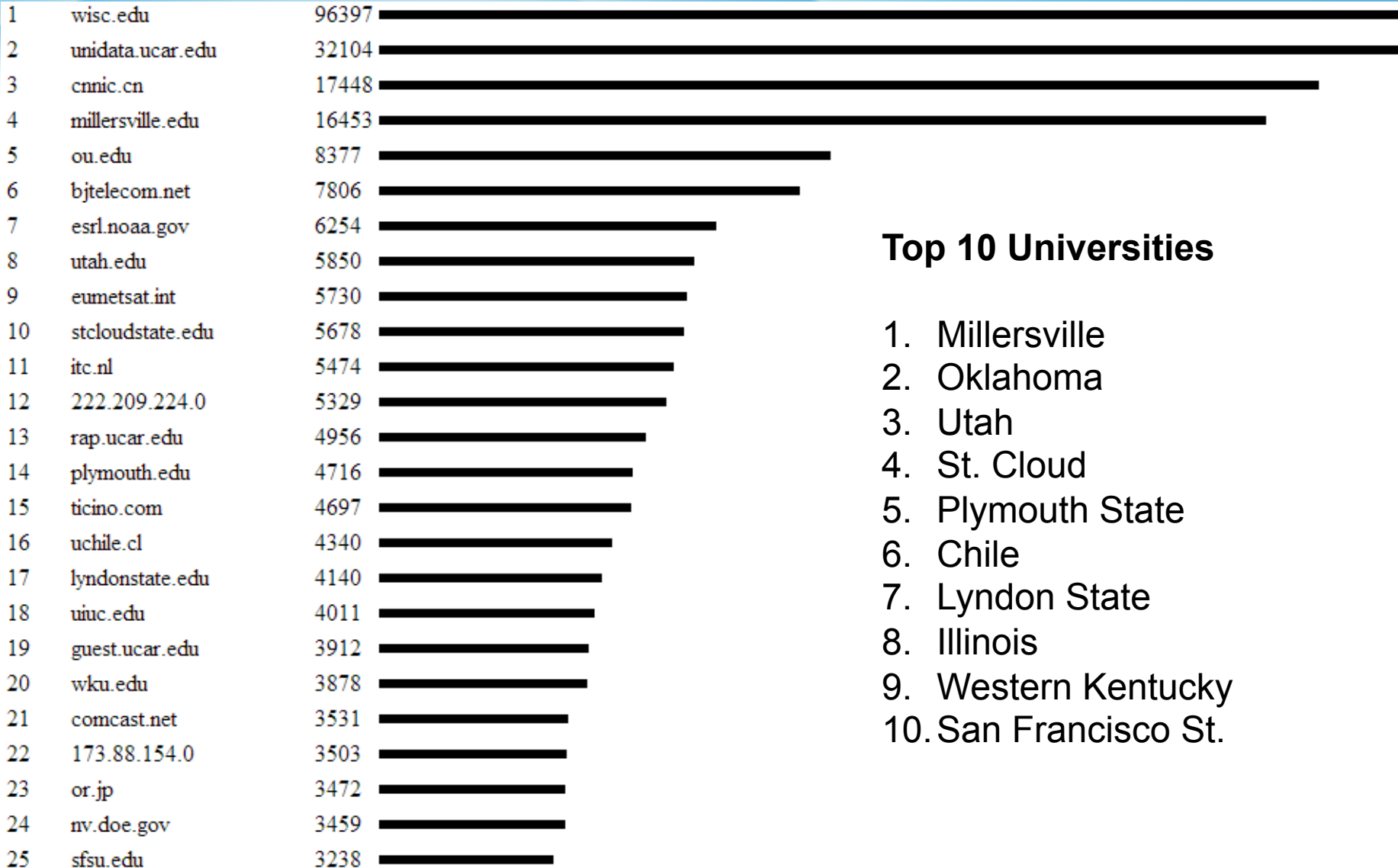
Integrated Data Viewer

❖ **IDV Milestones in the last six months.**

- the Java 7 / Java3D 1.6 transition.

- an early prototype of GEMPAK upper air format viewing capability.

- A well attended IDV Short Course during the AMS Annual Meeting.

- A number of Display advancements

  - New Image Chooser
  - Progressive Resolution capability
  - Matching Display Area

**The number of both casual and regular IDV users is increasing.**

IDV Usage

#	Domain	Count
1	wisc.edu	96397
2	unidata.ucar.edu	32104
3	cnnic.cn	17448
4	millersville.edu	16453
5	ou.edu	8377
6	bjtelecom.net	7806
7	esrl.noaa.gov	6254
8	utah.edu	5850
9	eumetsat.int	5730
10	stcloudstate.edu	5678
11	itc.nl	5474
12	222.209.224.0	5329
13	rap.ucar.edu	4956
14	plymouth.edu	4716
15	ticino.com	4697
16	uchile.cl	4340
17	lyndonstate.edu	4140
18	uiuc.edu	4011
19	guest.ucar.edu	3912
20	wku.edu	3878
21	comcast.net	3531
22	173.88.154.0	3503
23	or.jp	3472
24	nv.doe.gov	3459
25	sfsu.edu	3238

Top 10 Universities

1. Millersville
2. Oklahoma
3. Utah
4. St. Cloud
5. Plymouth State
6. Chile
7. Lyndon State
8. Illinois
9. Western Kentucky
10. San Francisco St.

GEMPAK

- GEMPAK7 was released in March 2014 and it incorporates a number of changes.

- NEX2GINI was re-tooled to better support generation of high-resolution NEXRAD composites such as Digital Hybrid Reflectivity, Digital Vertically Integrated Liquid, Enhanced Echo Tops, Hybrid Hydrometeor Classification.

- The national composite GINI images were made operational earlier this month.

- New program, **GPSCAT,** was added to plot scatterometer Significant Wave Height and
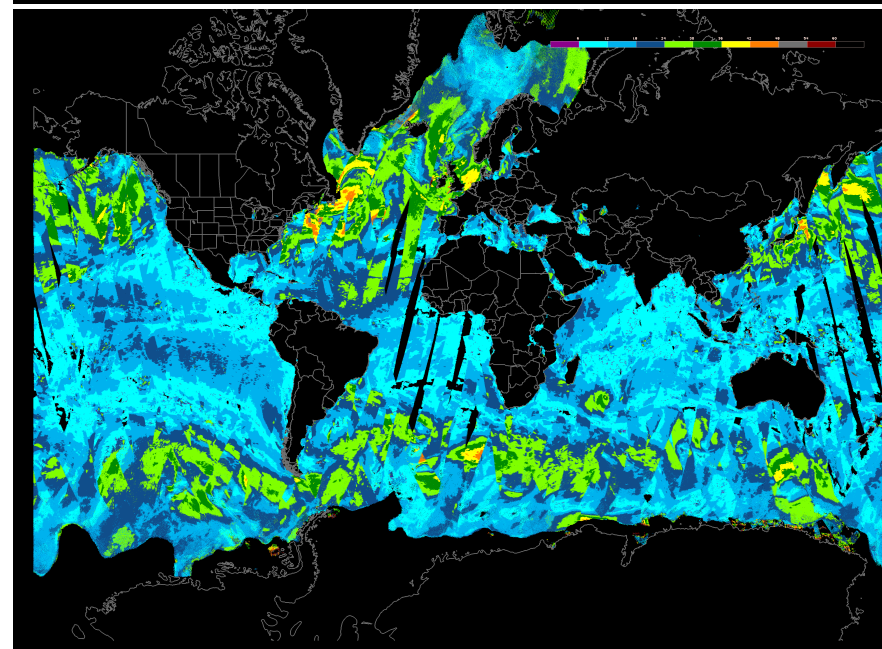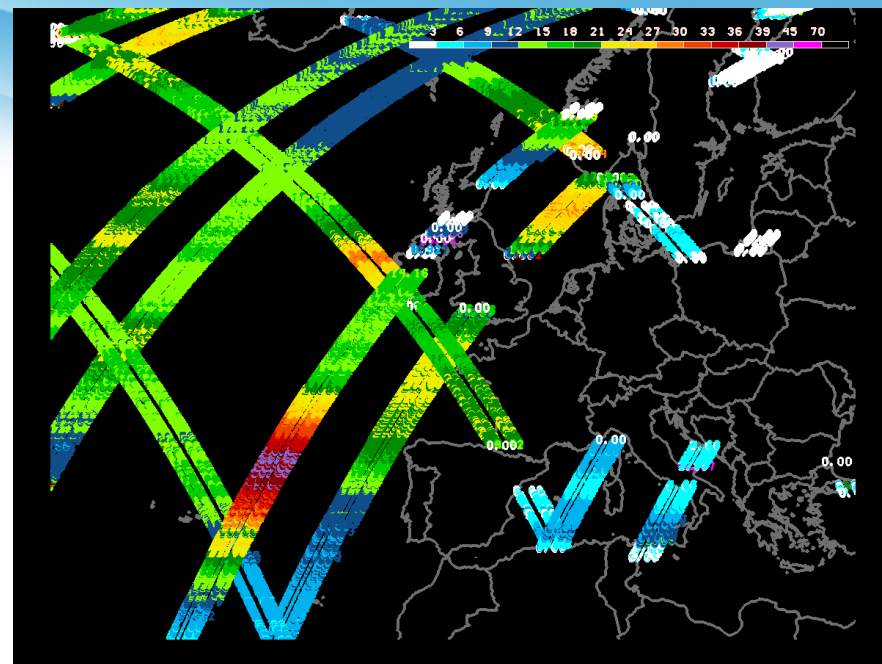
Rosetta

- The goal of Rosetta is to transform unstructured ASCII data into the netCDF format; once in this format, standard tools, such as the TDS, IDV, Python, and other analysis packages, can take advantage of these datasets with relative ease.

- Added the ability to publish converted files directly to RAMADDA and the ACADIS Gateway

- A live instance of Rosetta is hosted at Unidata for testing: https://rosetta.unidata.ucar.edu

- We need more sample datasets from you!

What would you like to do?

[+] Create a new template
[+] Upload, modify, and use an existing template
[−] Upload template and new data file, transform automatically

Rosetta

Python Efforts

- Our Python efforts will facilitate analysis of geoscience data by enabling data-proximate computations and analyses through IPython Notebook platform, which can be co-located with the data for analysis and visualization through web browsers.

- Our objective is to develop Python APIs tailored to Unidata technologies.

- **Initial efforts:** Conducting TDS-Python workshop, hosting netCDF-Python, incorporating Skew-T support in Matplotlib, and begun working on PyUDL, pyCWT, and pyCDM

- **NetCDF-Python:** Python interface to netCDF-4 was originally developed by Jeff Whitaker. Unidata has migrated it and hosting it on its GitHub site. **This has already encouraged new contributions and collaborations that are leveraging the efforts of other Python developers to improve and extend the software.**

- Unidata is continuing to be actively involved in EarthCube.

- Three of the five EarthCube Building Block proposals on which Unidata had partnered last year were funded.

- Unidata was again invited to participate on two new Building Block proposals. Both involve deployment of Unidata technologies in a cloud environment.

- Unidata is working with the EarthCube community to create a "Council of Data Facilities". I contributed to the writing of the charter for that body.

Unidata Instructional Videos on YouTube

* The Unidata You Tube channel now has 9 instructional videos.

* Brian Mapes, University of Miami, has contributed three instructional videos.

* The videos have been viewed more than 2000 times, a three-fold increase in the past six months.

AWIPS II

- Unidata is continuing to receive major releases of AWIPS II, including source code and Eclipse development environment, from the NWS.

- **The UPC has expanded the beta testing program to all U. S. universities, but there have been no new participants as yet.**

- AWIPS II version 14.2.1 is soon to be made available to Unidata. **It will support 64-bit Linux (RedHat and CentOS 6)**

- Standalone EDEX server at Unidata data is ingesting approximately 120 GB of data/day, including GFS and GEFS from CONDUIT and the full NEXRAD level 3 feed.

- Demonstrated AWIPS II at AMS Annual Meeting. WRF-NMM ingest and display proved successful at AMS 2014, and visualization support will be added to the next Unidata AWIPS II beta release.
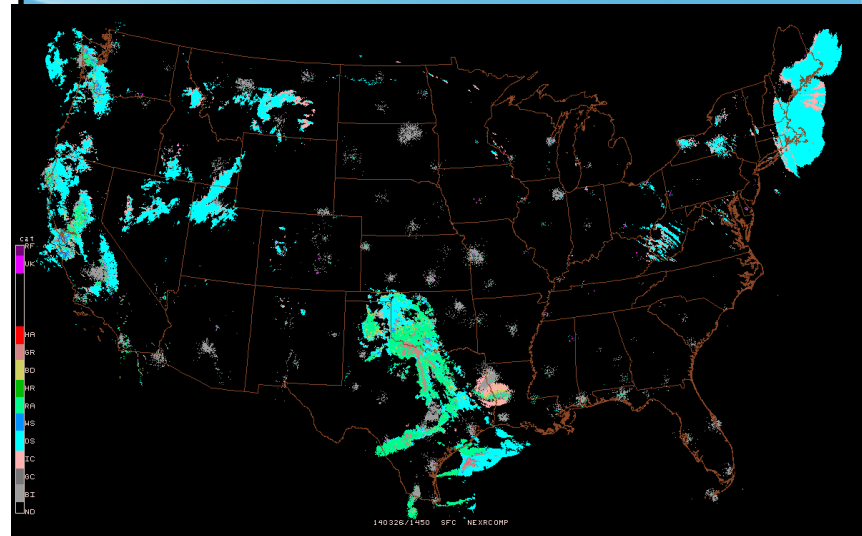
GINI to McIDAS Area file conversion for FNEXRAD composites has allowed the new high-res national images to be viewable as Area files in CAVE.

UCAR staff are discussing partnership opportunities with vendors that may bid on the next AWIPS II contract.

NOAA has already issued an RFI ahead of the recompetition.

2014 Community Equipment Awards

- Special consideration was given to proposals that focused on:

    – Installation of a prototype AWIPS II standalone EDEX server and CAVE client, coupled with the Unidata LDM, to test data ingest and display both locally, and using the CAVE thin client to connect to remote servers

    – Implementation of or pilot projects with remotely-accessible storage systems for geoscience data ("cloud-based storage")

    – Implementation of or pilot projects with remote server-based data analysis or visualization systems ("cloud-based analysis")

- **Unidata has received 13 proposals by the 14ᵗʰ March deadline.**

    – **There is a great deal of interest in AWIPS II deployment;**

Unidata Regional Workshop

• The University of Miami will be hosting and sponsoring a Unidata Regional Workshop on Friday and Saturday, April 18-19, 2014.

• The UPC staff will provide introductory and advanced training in the use of the IDV, RAMADDA, and other Unidata data and tools.

• The workshop will also provide an opportunity to present our plans for the next five years and seek feedback.

• So far, there are about 20 registrants from U. Miami, FIT, FIU, and U. South Florida.

2014 AMS Annual Meeting

- **Unidata organized a short course at the 2014 AMS Annual Meeting in Atlanta.**

- **Topic: "**Integrating WRF and Other Model Output with Remote and In-situ Observational Datasets using Unidata's Integrated Data Viewer"

- **There were 24 participants; the course was well received.**

We will also had a booth in the Exhibit area and a Table at the Career Fair.
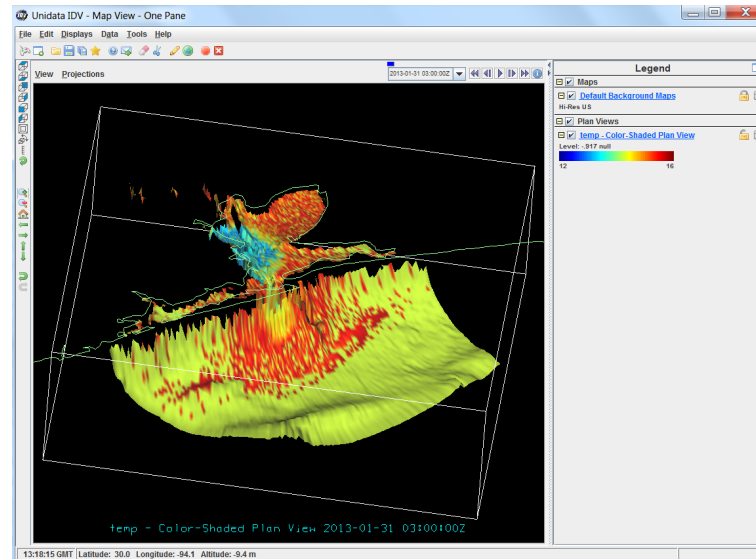
Russell L. DeSouza Award

And the winner is …

Rich Signell is an oceanographer at USGS Woods Hole Science Center in Massachusetts.

Rich has had a long-standing interest in data management, analysis, and visualization **and in promoting Unidata tools**.

He coauthored a paper on the benefits of using netCDF for ocean modeling more than 20 years ago!

State of the Program: A Snapshot

- Community relations   : Green
- Data flows   : Green
- Software development   : Yellow
- Support   : Green
- Collaborations & Outreach : Green
- Staffing   : Yellow
- Finances   : Red

Questions?