MEETING HOME

Agenda

Meeting Summary

Previous Meeting

Staff Status Reports

Users Committee Members

Strategic Advisory Committee Members

Acroynms

*Return to Committee Home*

# Unidata Joint Strategic Advisory and Users Committee Meeting

## 30 September - 2 October 2015

## Unidata Program Center Boulder, Colorado

## DRAFT Agenda

## Wednesday, 30 September

**New Committee Member Session (UPC main conference room, UCAR FL4)**

**10:00-12:00 - New Committee Member Welcome**

- Introduction to Unidata projects and team

**12:00-1:00 - Lunch**

**Users Committee Session (Center Green main auditorium, UCAR GG1)**

**1:00-1:30 - Welcome and Administrative Items**

- Introductions
- Review of Action Items
- Date for Spring meeting

**1:30-2:15 - Triennial Debrief**

- Lessons learned (Josh Young)
- Follow on actions
- Survey of participants (did it make a difference)

**2:15-2:45 - Discussion of Calls to Colleagues**

**2:45-3:00 - Break**

**3:00-4:00 - Planning Community Survey**

**4:00-4:30 - Discussion of Status Reports**

**4:30 - Retirement reception for Russ Rew (CG Patio)**

## Thursday, 1 October

**Joint Committee Session (Center Green main auditorium, UCAR GG1)**

**8:00-8:30 - Continental breakfast**

**8:30-9:00 - Welcome** - William Gallus

Welcome new members & introductions around the room

**9:00-9:30 - Users Committee report** - Kevin Tyle

**9:30-10:15 - Director's report** - Mohan Ramamurthy

**10:15-10:30 - Break**

**10:30-11:00 - AWIPS II** - Scott Jacobs

**11:00-11:30 - EarthCube** - Mohan Ramamurthy

**11:30-12:15 - DeSouza Award and Lecture** - Scott Jacobs

**12:15-1:15 - Lunch**

**1:15–1:30 - Group photo (Center Green patio)**

**1:30-2:00 - CONDUIT** - Scott Jacobs

**2:00-2:20 - Unidata Staff Retreat Update**

**2:20-2:40 - ADDIT Request for Partners** - Josh Young

**2:40-3:00 - UCAR Update** - Mohan Ramamurthy

**3:00-3:15 - Break**

**3:15-3:45 - Docker** - Ryan May

**3:45-4:15 - NOAA Big Data** – Mohan Ramamurthy

**4:15-4:45 - Unidata and Agency PARR Plans** - Josh Young

**4:45-5:00 - Miscellaneous discussion**

**5:00 Adjourn**

**6:00 Collaborative discussion on the day's proceedings over dinner at Rio Grande (map)**

## Friday, 2 October

**Strategic Advisory Committee Session (UPC main conference room, UCAR FL4)**

**8:00-8:30 - Continental Breakfast**

**8:30-9:00 - Administrative Items** - Bill Gallus

- Review of Action Items
- Date for Spring meeting

**9:00-9:30 - Budget Report** - Terry Mitchell-Sur

**9:30-10:00 - NASA Report** - Chris Lynnes

**10:00-10:15 - Break**

**10:15-10:45 - NOAA Report** - Jeff De La Beaujardiere

**10:45-11:15 - Next Unidata Program Center Strategic Plan**

**11:15-12:00 - UCAR Community Programs Discussion**

**12:00-12:30 - All Other Business**

**12:30 Adjourn**

# Status Report: ACADIS

*April - September 2015*

*Mohan Ramamurthy, Sean Arms, Jeff Weber*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *The ACADIS Data Portal is creating an effective way to access Arctic data*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *Unidata is creating an ASCII to netCDF translation tool that will allow a large amount of Arctic data to be translated to netCDF CF*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *ACADIS is an exemplar for data portals*

4. **Build, support, and advocate for the diverse geoscience community**
   *ACADIS continues to champion useful access to data holdings*

## Activities Since the Last Status Report

### The Rosetta project

Unidata's main contribution to the ACADIS project was the creation of Rosetta, a web-based data translator. The Rosetta project is now being developed under the THREDDS project umbrella. Rosetta will continue to be developed and maintained for the Unidata community, independent of the status of the NSF recompete proposal for the archiving and management of Arctic data.

## Ongoing Activities

The ACADIS team has submitted a proposal for the recompete of the archiving and management of Arctic data and is currently waiting for a decision from NSF.

Creation of visual displays of data transformed by Rosetta.

---

# Status Report: AWIPS II and GEMPAK

*April - September 2015*

*Michael James*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Both *AWIPS II and GEMPAK are freely available, and both incorporate LDM/IDD technology for efficiently accessing geoscience data.*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *Both AWIPS II and GEMPAK are open-source, and while GEMPAK is now in maintenance mode (not actively developed), AWIPS II is continuously being developed to support new data types and facilitate easier use.*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *Unidata is the only known entity to provide a freely-available and non-operational version of the AWIPS II software package.*

4. **Build, support, and advocate for the diverse geoscience community**
   *Using LDM/IDD technology to provide access to real-time meteorological data; providing visualization tools for data analysis.*

## Activities Since the Last Status Report

### AWIPS II

Unidata AWIPS II 14.4.1 was released at the beginning of September 2015, and includes significant changes to the system compared with the operational build, the focus being on minimizing the resource footprint and expanding the supported data types.

Some of the improvements and features in 14.4.1 include:

- **Server-side compression** to reduce data transfer sizes by an order of magnitude.
- Client-side **caching of maps and data** resources to reduce querying and loading time.
- Postgres, Pypies, Qpid and EDEX now start automatically on machine reboot.
- **Python Data Access Framework** allows users to write Python scripts to query remote EDEX maps and data.
- D2D single-pane default perspective, toggleable to 5-pane "classic" view.
- NCP resources can now be loaded directly from the D2D perspective, making available **GEMPAK-like grid display** without having to switch to the NCP

perspective.
- **Multi-Radar/Multi-Sensor** (MRMS) full support for decoding and displaying gzipped grib2 files from NSSL.
- Support for decoding and displaying **USPLN lightning data.**
- **NEXRAD3** menu includes full set of NEXRAD and TDWR sites, organized by site.
- **NEXRAD3 mosaic** capability renders a national composite of certain level 3 products queried from individual station files (functional but slow).
- **Save/Load Bundle** feature added for D2D, allowing users the option of saving bundles (as xml files) locally to disk or synced in ~/caveData to the EDEX server.
- **McIDAS AREA files** from the **UNIWISC** IDD feed decoded and displayed by the satellite plugins.
- NO CLIPPING OF GRIDS TO LOCALIZATION AREA
- Java 1.6 -> 1.7
- Fixed problem with how LDM writes **CMC-REG** grib1 messages which caused EDEX ingestGrib JVM to crash in 14.2.
- Full south polar stereographic **UNIWISC** McIDAS area file support.
- Partial **GVAR native projection** support.
- Various D2D UI improvement: max number of frames increased to 400, "Map Scale" menu now includes satellite-specific map projections (**Arctic, Antarctica, Alaska-Hawaii-PR Regional, East/West CONUS, Global, etc.**), Hydro/NCEP and Upper Air menus are now visible.
- Ingest JVMs now run with maximum 4096M of memory.
- Qpid components updated from 0.30 -> 0.32
- Various fixes to EDEX-side hydro applications.

## Ongoing Activities

- MRMS support requires the LDM use a modified version of the gempak “dcgunzip” decoder to decompress the grib2.gz files to disk, and then notify Qpid via a script “qpidNotify” (rather than edexBridge) that the product is available for ingest. Strangely, the operational build of AWIPS II has incomplete grib tables for MRMS processing, so the Unidata release is ahead of NWS on this matter. Still, the resource strain is too much for a standalone installation (with MRMS HDF5 files reaching 10 GB per hour, processing of NGRID and CONDUIT grids are slowed to the point of being unuseable), so a distributed EDEX environment (in the cloud, most likely) is needed to find a resource equilibrium.
- I have demonstrated the EDEX Python Data Access Framework (ufpy) to query and display gridded data (matplotlib) and upper air data (metpy), and work continues to provide more documentation and examples of how the framework can be used. The goal is to include ufpy as a means of requesting and displaying data inside in Unidata Siphon project, as a parallel to thredds/netCDF access.
- Further support of McIDAS GVAR native projection in EDEX, both on the processing side (currently seeing persistence errors when decoding multiple files) and on the visualization side (progressive resolution tiling is problematic with high-resolution UNIWISC files).
- (Exploratory) How best can netCDF-Java handle EDEX data? Is it possible to use existing (but currently hidden) EDEX technology to read from TDS?

### GEMPAK

GEMPAK v7.2.0 was released on August 4, 2015, incorporating NCEP table and map updates, as well as the following fixes and improvements:

- GFS 0.25 degree support
- Correction to allow decoding of 24hr min temp
- TDWR product fixes for nmap2 and gp/radar programs
- GEMPAK download access has been changed from registration-required to freely-available, and I have begun to maintain release packages (zip and tar.gz) on github for distribution rather than hosting on our own web server.

#### Ongoing Activities

- Maintain releases in parallel with NCEP operational table updates and fixes
- Incorporate a GUI-less GEMPAK build with the AWIPS II CAVE client which can access data from remote EDEX server.

## New Activities

#### Over the next three months, we plan to organize or take part in the following:

- Massively expand AWIPS II user documentation.

#### Over the next twelve months, we plan to organize or take part in the following:

- Python DAF incorporated into existing Unidata Python technologies
- Deploy a front-end web server (cherrypy is already installed with EDEX) for all EDEX machines to give users and administrators an interactive inventory of what data products are available.

## Relevant Metrics

Web server statistics?
Data download statistics?
Something else useful?

---

Prepared *September 2015*

# Status Report: Cloud Computing Activities

*April - September 2015*

*Sean Arms, Julien Chastang, Ethan Davis, Steve Emmerson, Ward Fisher, Tom Hollingshead, Michael James, Ryan May, Jennifer Oxelson, Mike Schmidt, Christian Ward-Garrison, Jeff Weber, Tom Yoksas*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *Making Unidata data streams available via various commercial and private cloud services will allow subscribers to those services to access data quickly and at low cost.*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *Running existing Unidata-developed and supported tools and processes (e.g. IDV, RAMADDA, generation of composite imagery) in a range of cloud environments makes these tools and data streams available to cloud service subscribers at low cost. It also gives us insight into how best to configure existing and new tools for most efficient use in these environments.*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *Unidata is uniquely positioned in our community to experiment with provision of both data and services in the cloud environment. Our efforts to determine the most efficient ways to make use of cloud resources will allow community members to forego at least some of the early, exploratory steps toward full use of cloud environments.*

4. **Build, support, and advocate for the diverse geoscience community**
   *[Build a bigger community]*

## Activities Since the Last Status Report

### Docker

With the goal of better serving our core community and in fulfillment of objectives articulated in Unidata 2018: __Transforming Geoscience through Innovative Data Services__ , Unidata is investigating how its technologies can best take advantage of cloud computing. To this end, we have been employing Docker container technology to streamline building, deploying, and running Unidata technology offerings in cloud-based resources. Specifically, we have created Docker images for the IDV, RAMADDA, THREDDS, Python with Unidata Technologies, and an initial attempt for the LDM, and we have been experimenting with these Docker containers in the Microsoft Azure and Amazon AWS commercial cloud computing environments.  Our preliminary efforts are available on Unidata's github repository. Also in one instance, we are using Docker technology operationally with the

testing of IDV bundles in the cloud.

While these efforts are promising initial steps, there are challenges ahead in making these technologies useful to our community. It is unlikely that most of our users will initially use these containers directly, rather they will be leveraged by experts on behalf of the community, or they will be abstracted from users by being integrated into a user-friendly workflow.

## AWS Training/Technical Discussions at the University of Wyoming

A number of Unidata technical staff traveled to Laramie, WY to meet with Amazon Web Services representatives for best practice training on the use of AWS resources including S3 and on efforts related to the NOAA Big Data Project. Meeting outside of Colorado was necessary to protect Amazon's Colorado sales tax position.

### Progress has been made on the following:

- Learning about Amazon's cloud infrastructure
- Designing an initial architecture to support putting all NEXRAD-2 data in Amazon's cloud
- Implementing a better NEXRAD-2 LDM decoder in Python for this cloud effort
- Implementation of a THREDDS Data Server on data stored in S3 on AWS

## Azure for Application Streaming/Unidata Service Hosting

Unidata has received a second year of Azure resources from Microsoft under the "Azure for Research" program. The primary focus of this award is continue work on creating an application-streaming platform for the IDV and other Unidata technologies. Secondary focus is on testing Unidata services in the Azure cloud, and examining the performance of Azure when hosting Docker instances.

With the release of Unidata AWIPS II 14.4.1 we have made available an EDEX Data Server in the Azure cloud (http://edex-cloud.unidata.ucar.edu:9581/services), and have set up a similar server privately for Embry-Riddle Aeronautical University on an Amazon EC-2 instance. Without a solid state drive these cloud deployments are ingesting a more limited data set than what can be ingested by a private EDEX Data Server located on campus. Bandwidth becomes an issue with very large data sets such as high-resolution gridded model HDF5 files, though the recent compression improvements to EDEX is shown to reduce data transfer rates by an order of magnitude.

## Progress has been made on the following:

- We have created a Dockerized version of the IDV bundled with a remote desktop/application streaming server. We are currently finishing up the first version of the associated web dashboard, "Cloud Control"
- We have deployed numerous services and instances to the Azure Cloud, mirroring

our experiments with the Amazon cloud infrastructure.
- We have learned a great deal about Microsoft's Azure cloud infrastructure.

# Ongoing Activities

**We plan to continue the following activities:**

- Use the LDM to move NEXRAD Level II data into AWS S3 buckets in real-time
- Develop enhanced procedures for recombining chunks of Level II data relayed in the IDD into full volume scans
- Develop TDS access to data stored in S3
- Deploy iPython notebooks that provide access to NEXRAD Level II data stored in S3

# New Activities

**Over the next three months, we plan to organize or take part in the following:**

- Deploy the first release of CloudIDV/Cloud Control to our community.
- Begin feeding data to Microsoft Azure for the Big Data Project

**Over the next twelve months, we plan to organize or take part in the following:**

- Implement machine images of our software for easy deployment in a virtual environment.
- Investigate containerizing as many Unidata services as possible.

**Beyond a one-year timeframe, we plan to organize or take part in the following:**

- Continue migrating Unidata services and software into the cloud, or cloud-suitable containers.

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

1. What clouds are our community using, either commercial or private?
2. What new cloud technologies are our community using/investigating on their own initiative?

---

# Status Report: Community Services

*April - September 2015*

*Doug Dirks, Jeff Weber, Joshua Young*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *We monitor and collaborate with data sources to stay apprised of impending changes and to advocate for the needs of our user community. We provide user workshops, tutorials, and community workshops to help build supportive relationships between community members.*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *We promote Unidata tools and software for multi-disciplinary use, with an eye toward finding additional research and educational communities that can benefit from our work.*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *We work with government and industry data providers to secure access to data for Unidata community members.*

4. **Build, support, and advocate for the diverse geoscience community**
   *We coordinate with our governing committees to find ways to expand Unidata's community participation. We use our web site, electronic newsletters, and social media to keep community members informed about enhanced data services, software tools, and cyberinfrastructure.*

   *We participate in UCAR/NCAR and NSF projects for underrepresented populations and minority communities (SOARS, AIHEC, outreach to HBCUs). We provide services and tools to facilitate education and research in diverse communities. We work to broaden the Unidata community by participating in student and professional conferences.*

## Activities Since the Last Status Report

### News@Unidata blog

Posts to the News@Unidata blog appear regularly, but not on a specific schedule. Some highlights:

- Unidata Workshop at the University of South Florida
- GIS Tutorials for the Atmospheric Sciences
- 2015 Community Equipment Awards
- MetPy: An Open Source Python Toolkit for Meteorology

- A New Look for Unidata's Web Site
- GFS 0.25° Model Output Added to CONDUIT
- 2015 Users Workshop Explores Python, Cloud Computing, Data Management
- DOIs Available for Unidata Technologies
- Unidata Intern Wraps Up Summer Projects
- Unidata Program Center Welcomes Tom Hollingshead
- Software release information
- Community job postings
- Community meetings and other announcements

**Dependencies, challenges, problems, and risks include:**

- Finding community members willing to contribute stories (or story ideas) for the blog is an ongoing challenge

## Community Outreach and Services

The community services group continues to actively reach out to and engage with Unidata community members.

**Progress has been made on the following:**

- Facilitating the release of GFS 0.25° model outputs to the community
- Organization of the 2015 Triennial
- Community engagement at upcoming professional society conferences (AGU, AMS, AGU Ocean Sciences)
- Identifying other resources for the DMRC
- Supported the launch of the National Weather Center with remote, daily weather briefings
- Engagement with CUAHSI to support the NFIE at the NWC
- Continue to serve on the CHUASI HIS standing committee
- Enhancing the Teaching Resource Network
- We continue to update Unidata's social media channels (Facebook, Twitter, Google+)
- We continue to publish short videos/screencasts on the Unidata YouTube channel.

**Dependencies, challenges, problems, and risks include:**

- Facilitating community adoption of new technological services (cloud, etc)
- Engagement with Unidata social media streams among community members is not particularly high.

# Ongoing Activities

**We plan to continue the following activities:**

- NAWIPS migration to AWIPS II, including the overall AWIPS II project

- Ongoing development of news articles for publication through News@Unidata
- Continue to support and contribute to governing committees
- Seminars
- Outreach
- Engagement with professional societies
- Support for cloud-related projects
- Further development of the Data Management Resource Center
- Further work on Agile Data Curation
- Site visits as the budget allows
- Continued additions to the Teaching Resource Network

# New Activities

**Over the next three months, we plan to organize or take part in the following:**

- Expanded emphasis on cloud-related activities
- Support the Users Committee in conducting a community survey

**Over the next twelve months, we plan to organize or take part in the following:**

- Coordinate a pilot effort to develop online training materials focused on Python and Unidata services and tools
- Consult with NOAA NWS partners regarding the outcomes of the NWS SOO survey
- Engage partners to develop case studies for the DMRC
- Engage community to contribute curriculum and data to the Teaching Resource Network

**Beyond a one-year timeframe, we plan to organize or take part in the following:**

- Conduct NWS SOO type surveys in other lines of NOAA or other Federal partners
- Provide additional data management and cloud-related training

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

1. Are you using cloud-based services currently or do you have any plans to do so?

2. We would like to document the use of various open-source services for data management and the user experience associated with those services. We are looking for various size projects (e.g 1 grad student, 1 PI, or a small project ~ 3 people). Do you have any suggested partnerships?

# Relevant Metrics

Statistics from the Community pages on the Unidata web site. Comparisons are made with statistics from the previous six-month period.

### All community pages

Most recent six months:

- 39,360 unique pageviews (40,128 in previous period)
- 6.9% of total unique pageviews to site (7.1% in previous period)

**Note:** This value is artificially low due to an interruption in collection of statistics for blog visits in June, July, and August. (Rough estimate of total unique pageviews: 54,000.)

### Top community pages

1. All blog pages
   19,574 unique pageviews (27,062 in previous period)
   50% of total community pageviews (67% in previous period)
   **Note:** Statistics collected for blog visits were interrupted after our switch to the new web server on June 16, 2015, which means that the blog pageviews count is missing 2.5 months worth of traffic. The actual traffic would have been significantly higher. (Rough estimate: 35,000 unique views.)
2. www.unidata.ucar.edu/events/
   12,010 unique pageviews (4676 in previous period)
   31% of total community pageviews (11.6% in previous period)
3. www.unidata.ucar.edu/about/
   3714 unique pageviews (3590 in previous period)
   9.4% of total community pageviews (9.0% in previous period)
4. www.unidata.ucar.edu/community/
   3164 unique pageviews (3825 in previous period)
   8.0% of total community pageviews (9.5% in previous period)

### Social media statistics, September 15, 2015

1. # of Twitter followers: 450 (up from 397)
2. # of Facebook followers: 428 (up from 383)

---

Prepared *September 2015*

# Status Report: Community Equipment Awards

**Sponsored by the National Science Foundation**

*April - September 2015*

*Admin Group*

The NSF provides the Unidata Program Center up to $100k in equipment grant funds each year. In alignment with the Unidata 2018 proposal, the Equipment Awards Program is designed to broaden participation and promote the use of Unidata tools and systems (e.g., THREDDS, NetCDF, IDV, GIS connections) to support education and research on various aspects of climate studies (e.g., diagnostics, change and impacts), by providing grants to be used in the procurement of new computers and equipment including upgrades to existing classroom and laboratory equipment.

This year, special consideration was given to proposals that included one or more of the following:

- 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")

A Request for Proposals was sent out on December 22, 2014 with a March 06, 2015 submission deadline. The Review Panel met on March 25 at the Unidata Program Center and recommended that the following proposals be awarded:

- Plymouth State University, Brendan Hoch, Transitioning to the IDV-CAVE: Improving Classroom Technology for Meteorology, Proposal
- University of California San Diego, Tom DeFanti, Flash I/0 Network Appliance (FIONA) connected to the 40Gb/s PRISM network at UC San Diego for worldwide access to the IDD, Proposal
- University of Nebraska Lincoln, Adam L. Houston, A Standalone EDEX Server and Enhanced Local IDD/LDM Infrastructure at the University of Nebraska - Lincoln, Proposal
- University of Wisconsin Milwaukee, A. Clark Evans, Deployment of AWIPS II at the University of Wisconsin-Milwaukee, Proposal
- Western Kentucky University, Joshua D. Durkee, Expanding Unidata Visualization and Data Analysis for Innovative Meteorological Education at Western Kentucky University, Proposal

Congratulations to all of the recipients and a special thank you to the Review Panel and the NSF for making the Equipment Awards program possible.

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

1. **Possible theme(s) for the 2016 Unidata Community Equipment Awards;**
2. **Volunteers to serve on the 2016 Review Panel**;
3. **Suggestions from previous panel members on how to improve the program**

# Relevant Metrics

Since taking over the management and administration of the Equipment Awards program in 2003 on behalf of the NSF, Unidata has made 82 awards totaling over $1,000,000.

---

Prepared  September 2015

# Status Report: IDV with RAMADDA

*April - September 2015*

*Yuan Ho, Julien Chastang*

This report updates the status of Unidata's Integrated Data Viewer (IDV) development efforts since the last report (April, 2015).

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   The IDV is a state of the art geoscience visualization application. It gives users the ability to view and analyze a rich set of geoscience data, including real time data, in a seamless and integrated fashion. This analysis is captured in IDV bundles. RAMADDA is a content management system and service specifically tailored towards the sharing and distribution of IDV bundles facilitating distribution of scientific data and analysis.

2. **Develop and provide open-source tools for effective use of geoscience data**
   The IDV has been an open-source project for several years. The IDV is available on the github version control platform for greater open-source collaboration. The IDV provides users the unparalleled ability to analyze, integrate, and visualize heterogeneous geoscience data in two, three, and four dimensions. The IDV coupled with RAMADDA enables geoscience specialists the capability to share and collaborate their IDV analysis via social scientific networks.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   RAMADDA allows geoscience specialists the ability to search and publish their IDV bundles on-line. Unidata's RAMADDA installation enables the IDV team to communicate more effectively to our users concerning their IDV issues. Specifically, during support ticket conversations, the IDV team requests that users upload pertinent data to RAMADDA for analysis. The IDV team also takes advantage of RAMADDA to share instructional IDV screencasts with users.

4. **Build, support, and advocate for the diverse geoscience community**
   Unidata offers yearly multi-day training and occasionally regional workshops for IDV and RAMADDA. The IDV coupled with RAMADDA enables our earth science community partners to distribute geoscience data and metadata through web-based technologies thereby fostering scientific collaborations. Moreover, the IDV's ability to share bundles through RAMADDA creates a scientific social and collaborative network for the geoscience community.

# Activities Since the Last Status Report

## IDV System Changes

__Java Version 1.8__

The IDV is now fully running on Java 8. Specifically, the IDV is being distributed with and running under Java 8 JVM. The IDV is also now in a position to make use of Java 8 language features.

__Jython Version 2.7__

Updated the Jython library used by the IDV to version 2.7.

__Digital Object Identifier__

For improved citation purposes, the IDV now has a Digital Object Identifier (DOI): doi:10.5065/D6RN35XM.

__Latest netCDF-Java Version__

The version of the netCDF-Java library currently distributed with the IDV is the 4.6.2. Its main feature is a rewrite of GRIB handling. This version of netCDF-Java includes changes: AWIPSsat projection, WRF lat/lon coordinates with time dependence, Updated grib1 table for NLDAS, handling of scalar runtime coordinate for GRIB collections. See the netCDF-Java Library for a more details on these changes.

## IDV Display Changes

__Forecast hour Capability __

The IDV can now calculate the forecast hours based on the run time coordinate system. These  values can be accessed through the forecast hour macro in the display properties window.

__New Grid Trajectory and Curly Vector Capabilities__

Improvements in the  2D grid trajectory and the curly vector display features are based on enhancements in the VisAD library. Both of these displays are rendered as flow vectors. The vector plan view display is created first. At this point, there are radio buttons to switch to the trajectory or curly vector display. The trajectory and the curly vector can also be colored by the wind speed.

__3D Adaptive Resolution__

The Adaptive Resolution (AR) feature in the IDV has been expanded to cover several three-dimensional displays. When creating a three-dimensional or cross section display, the IDV calculates the resolution of the map view window, dynamically sets the data sampling in all 3 dimensions, and loads sufficient data to generate the image.

__Support for CF Trajectory Format__

The IDV now supports the reading and display of trajectory data as described in H.4. Trajectory Data of the CF conventions.

__VisAD Vertical Scale Labeling__

Incorporated improved vertical scale labeling available in from VisAD.

## New Testing Suite

The IDV testing suite was rewritten. In particular,

- The outdated testing bundles were fixed with up-to-date variable names and data servers
- The testing suite was dockerized to make it portable
- A crude image comparison algorithm was developed to show the most problematic bundles first
- This dockerized testing suite was deployed to http://unidata-idvtest.cloudapp.net/compare.html (link not working, but will be by usercomm hopefully)

and containerized with Docker to make it portable (which was a feature that was lacking in the past). The testing suite

## IDV Release

The IDV Team released version 5.2 of the IDV.

## IDV Proposal Work

__NASA ROSES__

With our University of Wisconsin, Space Science and Engineering Center partners, we submitted a proposal to NASA entitled, “Interactive Algorithm Development and Product Validation through Innovative Data Access and Visualization Methods”.

## IDV Publication Highlights

The IDV was used for the generation of figures in the Nature publication: Relativistic electron avalanches as a thunderstorm discharge competing with lightning.

## RAMADDA

Docker is a new cloud-centric technology that borrows from the notion of containers from the shipping industry to facilitate installation and deployment of server side applications. We have implemented a prototype Docker container for easy distribution and installation of RAMADDA in a cloud environment. We aim to couple this effort with a Dockerized LDM with the goal of experimenting serving data in a cloud environment.

## IDV and RAMADDA Training and Conference Attendance

__IDV Training in San Jose, Costa Rica__

Unidata conducted an IDV/McIDAS-V/RAMADDA training workshop in San Jose, Costa Rica during the week of May 4, 2015. This effort is being organized by Maria del Rosario Alfaro Ocampo, a UCAR/JOSS employee who works in the International Activities Office at the NWS in Washington, DC, in support of National Weather Services of Central America.

The training will focus on integrating WRF output with GOES satellite imagery in support of generation of timely weather warnings. Unidata's participation in this workshop is being funded by USAID.

__IDV Training Workshop, August, 2015__

Conducted one-on-one IDV training centered on WRF output with a user from France's Direction générale de l'armement (DGA) and a user from the Mesoscale and Microscale Meteorology (MMM) Laboratory at NCAR.

__IDV Instructional Videos__

Produced a new IDV training video entitled Exploring Earthquake CSV Data.

---

# Ongoing Activities

**We plan to continue the following activities:**

__IDV in the Cloud __

With the goal of better serving our core community and in fulfillment of objectives articulated in “Unidata 2018: Transforming Geoscience through Innovative Data Services”, the IDV team will continue to investigate how its technologies can best take advantage of cloud computing. To this end, we have been employing Docker container technology to streamline building, deploying, and running Unidata technology offerings in cloud- based resources. Specifically, we have created Docker images for the IDV, RAMADDA, we are working on the LDM which, coupled with RAMADDA, will allow for the serving of real-time data in a cloud environment for IDV users. We have been experimenting with these Docker containers in the Microsoft Azure and Amazon.

__IDV Instructional Videos __

We plan to continue producing more instructional videos on the IDV. We would appreciate input and suggestions on specific video topics.

# New Activities

**Over the few months, we plan to organize or take part in the following:**

__2016 American Meteorological Society's Annual Meeting in La Nouvelle-Orléans __

- We are preparing to demonstrate the IDV and RAMADDA at the 2016 AMS annual meeting.
- We submitted an abstract for the AMS 2016 Environmental Information Processing Technologies (EIPT) entitled “UniCloud: Docker Use at Unidata”.
- We aim to demonstrate RAMADDA running in the cloud serving a real-time data feed.

# Areas for Committee Feedback

We have no questions at this time.

# Relevant Metrics

The IDV team continues to provide the geoscience community with high-quality support through e-support software and idv-users maillist. The volume of e-support remains high and constitutes a large fraction of our daily activities. In the last half year the IDV team has closed ~100 e-support tickets. Also, we have reached an import e-support milestone. The

IDV team recently closed its 4,000th e-support ticket. Each individual ticket may and often does involve many back-and-forth messages.

Raw IDV usage metrics, are available here http://www.unidata.ucar.edu/software/idv/logging/left.html.

In the area of greater collaborative development, since the migration of the IDV project to github, we have closed a total of 85 “pull requests” or code contributions from internal and external collaborators.

In the area of online IDV training, the Youtube IDV instructional videos have been viewed over 6,115 times compared with 4,840 from six months ago. Mathew Dewey’s synoptic meteorology video has generated an especially great interest being viewed 415 times.

---

Prepared September 2015

# Status Report: International Activities

*April - September 2015*

*Tom Yoksas*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Develop and provide open-source tools for effective use of geoscience data**
   *The majority of tools downloadable from Unidata are available free-of-charge to everyone (the exception being McIDAS-X).*

2. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *Activities of the Unidata Program Center are routinely provided to the worldwide atmospheric science community. Strategic partnerships with leading organizations in other countries minimize the impact on UPC staff.*

3. **Build, support, and advocate for the diverse geoscience community**
   *By informing the international atmospheric science community of the products, data and services available in the Unidata Program, an extended community has been enabled.*
   *Non-U.S. users of products available from Unidata reflect, in a number of cases, minority constituencies in the U.S. atmospheric science community.*

## Activities Since the Last Status Report

### Regional Training Workshop

Unidata conducted a training workshop focused on using the IDV, McIDAS-V (for Hydra) and RAMADDA in San Jose, Costa Rica in May, 2015.

The training focused on integrating WRF output with GOES satellite imagery in support of generation of timely weather warnings. Unidata's participation in this workshop was funded by USAID.

## Ongoing Activities

- Data from UCAR GOES East/West ingest systems continue to be routinely accessed by international users in North, Central and South America using McIDAS-X, IDV, and McIDAS-V.

- Use of Unidata tools, especially netCDF, the IDV and GEMPAK, continues to grow

# Status Report: Internet Data Distribution

*April - September 2015*

*Mike Schmidt, Jeff Weber, Tom Yoksas*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *A project like the IDD demonstrates how sites can employ the LDM to move data in their own environments.*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *The IDD is powered by the Unidata LDM-6 which is made freely available to all. The Unidata NOAAPort ingest package is being used by a variety of university and non-university community members. Both the LDM and NOAAPort ingest packages are being bundled by Raytheon in AWIPS-II.*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *The community-driven IDDs provide push data services to users an ever increasing community of global educators and researchers.*

4. **Build, support, and advocate for the diverse geoscience community**
   *Providing access to data in real-time is a fundamental Unidata activity.*
   *The IDD-Brasil, the South American peer of the North American IDD operated by the UPC, is helping to extend real-time data delivery outside of the U.S. to countries in South America and Africa. The Universidad de Costa Rica is experimenting with relaying data received in the IDD to Colombia.*

## Activities Since the Last Status Report

### Internet Data Distribution (IDD)

After an extensive evaluation period,  0.25 degree GFS data (which became operational in NCEP on January 14, 2015) was added to the CONDUIT data stream starting with the 12Z run on July, 28 . Monitoring has shown that peak CONDUIT data volumes increased from about 8 GB/hr to about 21 GB/hr for all forecast hours for the 0.25 degree GFS.

The increase in aggregate data volume that results from the addition of the 0.25 degree GFS and HRRR data from NOAA/GSD can be seen by comparing the volume on our IDD test leaf node, lead.unidata.ucar.edu with that on one of the idd.unidata.ucar.edu real server backends shown below:

``bqb
Data Volume Summary for **lead.unidata.ucar.edu**

Maximum hourly volume  63172.218 M bytes/hour
Average hourly volume  31509.536 M bytes/hour

Average products per hour     359719 prods/hour

| Feed | Average (M byte/hour) | | Maximum (M byte/hour) | Products number/hour |
|---|---|---|---|---|
| CONDUIT | 7543.544 | [ 23.941%] | 21413.167 | 83618.087 |
| FSL2 | 7494.892 | [ 23.786%] | 32739.626 | 12676.435 |
| NEXRAD2 | 5674.899 | [ 18.010%] | 8061.093 | 63293.500 |
| NGRID | 4687.849 | [ 14.878%] | 8107.771 | 33423.978 |
| NOTHER | 2118.776 | [ 6.724%] | 4479.228 | 6299.522 |
| NEXRAD3 | 1939.169 | [ 6.154%] | 2602.765 | 92942.761 |
| FNMOC | 1156.486 | [ 3.670%] | 3860.758 | 3251.283 |
| HDS | 355.593 | [ 1.129%] | 654.965 | 19128.804 |
| NIMAGE | 159.260 | [ 0.505%] | 263.313 | 202.565 |
| FNEXRAD | 128.916 | [ 0.409%] | 169.548 | 105.239 |
| GEM | 78.549 | [ 0.249%] | 495.761 | 757.891 |
| UNIWISC | 69.489 | [ 0.221%] | 117.668 | 43.783 |
| IDS\|DDPLUS | 62.497 | [ 0.198%] | 75.362 | 43338.783 |
| EXP | 36.210 | [ 0.115%] | 73.385 | 285.848 |
| LIGHTNING | 3.296 | [ 0.010%] | 6.446 | 349.196 |
| GPS | 0.111 | [ 0.000%] | 1.290 | 1.022 |

Data Volume Summary for **uni16.unidata.ucar.edu**

Maximum hourly volume  33304.768 M bytes/hour
Average hourly volume  19089.235 M bytes/hour

Average products per hour     292558 prods/hour

| Feed | Average (M byte/hour) | | Maximum (M byte/hour) | Products number/hour |
|---|---|---|---|---|
| NEXRAD2 | 5676.682 | [ 29.738%] | 8061.093 | 63316.913 |
| NGRID | 4691.342 | [ 24.576%] | 8107.771 | 33447.370 |
| CONDUIT | 2624.859 | [ 13.750%] | 10488.076 | 28922.630 |
| NOTHER | 2169.096 | [ 11.363%] | 4479.228 | 6431.217 |
| NEXRAD3 | 1940.247 | [ 10.164%] | 2602.765 | 93001.304 |
| FNMOC | 1156.486 | [ 6.058%] | 3860.758 | 3251.283 |
| HDS | 355.714 | [ 1.863%] | 650.487 | 19136.674 |
| NIMAGE | 155.868 | [ 0.817%] | 263.313 | 200.348 |
| FNEXRAD | 94.550 | [ 0.495%] | 155.292 | 66.957 |
| GEM | 78.549 | [ 0.411%] | 495.761 | 757.891 |
| IDS\|DDPLUS | 62.529 | [ 0.328%] | 75.362 | 43366.174 |
| UNIWISC | 43.712 | [ 0.229%] | 106.520 | 25.783 |
| EXP | 36.213 | [ 0.190%] | 73.385 | 285.870 |
| LIGHTNING | 3.275 | [ 0.017%] | 6.446 | 346.413 |
| GPS | 0.112 | [ 0.001%] | 1.290 | 1.043 |

``bqe

Recently, top level IDD relays and the sites that they are feeding CONDUIT data to have been experiencing unusually high latencies that correspond with the transmission of the 0.25 degree GFS data. Current testing suggests that a large fraction of the latencies being experienced originate at or near NCEP. Investigations are ongoing.

# Ongoing Activities

**We plan to continue the following activities:**

- Unidata continues to receive High Resolution Rapid Refresh (**HRRR**) grids (both 2D and 3D fields) in an LDM/IDD feed from NOAA/GSD and feed these products to a small number (6) of university sites on **hrrr.unidata.ucar.edu**. Since HRRR and ESTOFS data were added to the NOAAPort Satellite Broadcast Network (SBN) in late September, 2014, continuing to relay the HRRR ingested from NOAA/GSD is considered to be of lesser importance and will be discontinued if the sites currently receiving the NOAA/GSD data are amenable.

- The HRRR is being experimentally served at:
  http://thredds-jumbo.unidata.ucar.edu/thredds/modelsHrrr.html
  (.xml for machines)

- Other data sets we are actively exploring with NOAA/GSD/ESRL are:
  - FIM
  - HIWPP

- HRRR and ESTOFS products were added to NOAAPort in late September, 2014. The following TINs announced these additions:

  http://www.nws.noaa.gov/os/notification/tin14-28hrrr-cca.htm
  http://www.nws.noaa.gov/os/notification/tin13-43estofs_noaaport_aaa.htm

  Briefly, these additions are comprised of:
  - HRRR: 81 products, hourly F00-15 each hour. CONUS 2.5km grid184. ~44 GB/day
  - ESTOFS: 3 products, hourly F00-F180, 00, 06, 12, 18z runs. CONUS 2.5km grid, Puerto Rico 1.25 km grid. ~2 GB/day

- HRRR fields and forecasts times that are not included in the NOAAPort expansion will be evaluated as additions to the CONDUIT IDD datastream.

- The UPC continues to relay FNMOC and CMC data model output directly to the community. FNMOC provides the COAMPS and NAVGEM model output and the CMC provides the GEM model output. Unidata has provided access to these data for the past 8 years, but on a "point-to-point" basis. GEM model output was converted from GRIB1 to GRIB2 in January. The CMC is now relaying output of there new

hi-resolution (15 km) GEM model to Unidata.

# NOAAPort Data Ingest

- The NOAAPort SBN, which transitioned from DVB-S to DVB-S2 in April/May 2011, was upgraded to support much higher throughput in August, 2014. Ingestion of the broadcast has been “operational” at the UPC since the upgrade.

- Comparison of our ingest metrics with other sites running our software (e.g., UW/SSEC, NOAA/GSD, LSU/SRCC, and a Northrop Grumman office in Northern Virginia) strongly suggested that signal quality was a major contributing factor in the problems were being experienced. **Considerable** effort to understand the data ingest problems being experienced was expended in the UPC. Experimentation demonstrated that that use of an older version of firmware on out Novra S300N receivers (V2R7 the version recommended by the manufacturer for our hardware) would produce errors in the UDP output UDP when the S300Ns were interrogated for status information, and this effect was, in turn, a function of signal quality. Ingest problems were verified by one of the commercial vendors of NOAAPort receipt systems who worked with Novra to correct this and other problems being experienced by S300N receivers.

  The situation of routinely experiencing high numbers of missed frames has largely been mitigated through a combination of hardware upgrades and by a Novra firmware upgrade that was aimed at dealing with the “small” packets routinely seen in the GOES product channel.

- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.12.14.

- Raytheon bundles a version LDM-6 with AWIPS-II and is actively using Unidata's NOAAPort ingest code at a variety of NOAA offices. Raytheon has provided code modifications and GRIB table updates needed to support new data to be added to in the NOAAPort expansion the UPC.

# Relevant Metrics

- Approximately **550** machines at **252** sites are running LDM-6 **and** reporting real time statistics to the UPC. Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

- A number of organizations/projects continue use the LDM to move substantial amounts of data that do not report statistics to Unidata: NOAA, NASA, USGS, USACE, Governments of Spain, South Korea, private companies, etc.).

- UCAR IDD toplevel relay node, **idd.unidata.ucar.edu**

The IDD relay cluster, described in the June 2005 CommunitE-letter article Unidata's IDD Cluster, routinely relays data to more  than 1250 downstream connections.

Data input to the cluster nodes averages around 20 GB/hr (~0.5 TB/day);  average data output from the entire cluster exceeds 2.9 Gbps (~32 TB/day); peak rates routinely exceed 6.4 Gbps (which would be ~70 TB/day if the rate was sustained).

Cluster real server backends and accumulator nodes routinely have instantaneous output volumes that exceed a Gpbs.  Bonding of pairs of Ethernet interfaces was needed to be able to support these output data rates.  The next generation of cluster machines will need to have 10 Gbps Ethernet capability.

The increase in data volume over the past six months is attributable to the addition of 0.25 degree GFS data to CONDUIT, the overall increase in the volume of data being transmitted in NOAAPort (which now routinely exceeds 10 GB/hr), and the increase in dual polarization NEXRAD data.  During the end of August/beginning of September GOES-R test period, the NOTHER datastream's pushed the total volume of data being sent over NOAAPort to peaks in excess of 20 GB/hr.

---

Prepared *September, 2015*

internationally.
- IDD-Brazil continues to deliver data via the LDM in Africa.

---

Prepared *September, 2015*

# Status Report: LDM

*April - September 2015*

*Steve Emmerson, Tom Yoksas, Mike Schmidt*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *By enabling researchers, teachers, and students to obtain a wide variety of meteorological and related data in near real time and at no cost via the Internet.*

2. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *By using the LDM to move data into the cloud.*

## Activities Since the Last Status Report

### LDM

The LDM is the primary software package by which research and education institutions obtain near real time meteorological and related data.

#### Progress has been made on the following:

- Bugs dealing with the last successfully-received data-product and the last successfully-processed data-product were found and squashed.
- Static code analysis by Coverity Scan was added to the LDM development pipeline.
- The product-queue was hardened against simultaneous access by multiple threads.
- The build procedure was improved by making it more general and robust.
- Versions 6.12.8 through 6.12.14 were released.

#### Dependencies, challenges, problems, and risks include:

- Dealing with the ever-increasing volume of data.
- Very low bus factor (if Emmerson gets hit by a bus, the LDM's in trouble).

### Multicast LDM (alias LDM-7)

The multicast LDM project is separately funded by CISE in NSF. The goal is to reduce the outgoing bandwidth requirement of the LDM -- yet retain the current level of reliability -- by converting it into a hybrid system that combines use of the new, semi-reliable multicast protocol developed at the University of Virginia with the time-tested unicast capability of

the current LDM.

**Progress has been made on the following:**

- Baseline statistics of LDM-6 behavior have been gathered.
- Testing of LDM-7 in a wide-area network has begun.

**Dependencies, challenges, problems, and risks include:**

- Time.
- Very low bus factor.

# Ongoing Activities

**We plan to continue the following activities:**

- Support and maintenance of the LDM
- Deployment and testing of LDM-7

# New Activities

**Over the next three months, we plan to organize or take part in the following:**

- Use of the LDM to bring data into the Cloud as part of the NOAA Big Data Project.

**Over the next twelve months, we plan to organize or take part in the following:**

- Use of the LDM to bring data into the Cloud as part of the NOAA Big data Project.

# Relevant Metrics

The LDM powers the IDD, whose statistics can be found here.

---

# Status Report: McIDAS

*April - September 2015*

*Tom Yoksas*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *McIDAS remains **the** application of choice for the satellite meteorology community. The Advanced Data Distribution Environment (ADDE) component of McIDAS was the first application offered by Unidata to provide remote, programmatic access to a wide variety of data that is important to the atmospheric science community.*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *The fifth generation of McIDAS, McIDAS-V, unlike its predecessors, is a fully open source application that is is in wide scale and growing use in the satellite meteorology community.*
   *McIDAS ADDE continues to evolve and provide access to increasing volumes of image and non-image data.*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *Concepts articulated in ADDE inspired the development of THREDDS (to address the lack of rich metadata available in ADDE) and RAMADDA. ADDE remains one of the most used data services in the Unidata suite serving over 4.5 TB per month. ADDE servers in the SSEC Data Center are currently serving over 1 TB per day.*

4. **Build, support, and advocate for the diverse geoscience community**
   *McIDAS is sought for use by those interested in satellite meteorology worldwide.*

## Activities Since the Last Status Report

Unidata McIDAS-X/XCD is in mostly in maintenance mode. Aside from routine updates/bugfixes to existing code and tables, the main thrust of development is to add indexing to ADDE datasets to speed access into large datasets.

### Prior Activities

- Unidata McIDAS version 2015 was made available in mid-summer. v2015 includes all SSEC versions up to and including the current release, v2015.1 and Unidata updates and bugfixes.

  Changes to Unidata McIDAS continue to be made through an **addendum** process.

**Progress has been made on the following:**

- Enhancement of the METEOSAT Second Generation (MSG) ADDE data server
- Introduction of an ADDE server for Himawari imagery (this is a precursor for a GOES-R ADDE server

# Ongoing Activities

**We plan to continue the following activities:**

- SSEC McIDAS Advisory Committee (MAC)

  The UPC (Yoksas, Ho) continues to participate as the Unidata representative to the McIDAS Advisory Committee (MAC) that is operated by SSEC.

  The MAC was assembled by UW/SSEC to advise SSEC on McIDAS-X users needs/concerns/desires for development in the next generation McIDAS, McIDAS-V. The MAC was modeled after the Unidata IDV Steering Committee.

- Interest in McIDAS by Non-core Users

  The UPC continues to receive requests for McIDAS from international university users, U.S. government agencies and other non-traditional Unidata users (e.g., private businesses, etc.). Government agencies and non-traditional Unidata users are referred to UW/SSEC for access to McIDAS; international educational community user requests are granted on a case-by-case basis after they provide a clear statement of their acceptance of the terms of use provided by SSEC.

# New Activities

## Ongoing Activities

Continued support of existing and new community members.

## New Activities

Implement indexing for ADDE image datasets to speed up access especially in large and archive datasets. A preliminary design for ADDE image dataset indexing has been made. Investigations for how to integrate the new capabilities into the suite of existing ADDE servers is in progress.

Add support for new types of data when they become available, otherwise McIDAS-X support is in maintenance mode.

# Relevant Metrics

- Internet2 (I2) bandwidth usage by the McIDAS ADDE protocol routinely exceeds 8 TB/week.
- McIDAS Inquiry Metrics

# ldm-mcidas Decoders Activities

### Development

**ldm-mcidas** releases are made when needed to support changes in software development and operating system environments. **ldm-mcidas** v2012 was released at the end of September, 2012. This release addressed building on newer OS versions.

# Geostationary Satellite Data Ingest

Unidata continues to ingest GOES-East and GOES-West imager data at the UCAR Foothills Lab campus in Boulder.

- Direct, programmatic access to real-time GOES-East (GOES-13) data via McIDAS ADDE results in an average of approx. 2 TB/month.
- Direct, programmatic access to real-time GOES-West (GOES-15) data via McIDAS ADDE results in an average of approx.1.7 TB/month.

# Planned Activities

### Ongoing Activities

Continued ingest and serving of GOES-East and GOES-West imagery from the current constellation of GOES GVAR platforms. This effort requires sporadic maintenance of the satellite ingest and data serving equipment.

### New/future Activities

Install a GOES-R downlink, processing and data serving capability at the NCAR Mesa Lab using the easternmost of the old USAN satellite pads. NOAA GOES-R office funding for this activity is anticipated to become available in the beginning of January, 2016. Launch of GOES-R is currently planned for fall, 2016.

Investigate the feasibility of moving the GOES-R imagery and products to “the cloud” in real-time. Preliminary discussions with Amazon Web Services representatives have already take place, and they are very interested in the GOES-R data being made available in the same ways as Landsat imagery.

Prepared September, 2015

# Status Report: netCDF

*April - September 2015*

*Ward Fisher, Dennis Heimbigner, Russ Rew*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *by developing netCDF and related cyberinfrastructure solutions to facilitate local and remote access to scientific data.*

2. **Develop and provide open-source tools for effective use of geoscience data**
   *by supporting use of netCDF and related technologies for analyzing, integrating, and visualizing multidimensional geoscience data; enabling effective use of very large data sets; and accessing, managing, and sharing collections of heterogeneous data from diverse sources.*

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *by developing useful data models, frameworks, and protocols for geoscience data; advancing geoscience data and metadata standards and conventions; and providing information and guidance on emerging cyberinfrastructure trends and technologies.*

4. **Build, support, and advocate for the diverse geoscience community**
   *by providing expertise in implementing effective data management, conducting training workshops, responding to support questions, maintaining comprehensive documentation, maintaining example programs and files, and keeping online FAQs, best practices, and web site up to date; fostering interactions between community members; and advocating community perspectives at scientific meetings, conferences, and other venues.*

## Activities Since the Last Status Report

We are using JIRA, GitHub tools for C, Fortran and C++ interfaces to provide transparent feature development, handle performance issues, fix bugs, deploy new releases and to collaborate with other developers. Additionally, we are using docker technology to run netCDF-C, Fortran and C++ regression and continuous integration tests. We currently have 98 open issues for netCDF-C, 18 open issues for netCDF-Fortran, and 3 open issues for netCDF-C++. The netCDF Java interface is maintained by the Unidata CDM/TDS group (which also uses Jira and GitHub), and we collaborate with external developers to maintain the netCDF Python interface.

- In the netCDF group, progress has been made in the following areas since the last status report:
  - We presented on netCDF at the Unidata Python training workshop in July.
  - We were invited speakers at the Rocky Mountain Advanced Computing

Consortium (RMACC) HPC Conference in August, speaking on netCDF.
    - We have incorporated support for the 64-bit-everything netCDF format from the parallel netCDF project at Argonne and Northwestern into a branch of the netCDF project, pending inclusion in the next release.
    - Further enhancements to the netCDF documentation.
    - Extended continuous integration platforms.
    - Collaborated with HDF5 and Engility Corp on a proposal for a DoD-funded project to incorporate new HPC functionality (ExaHDF5) into netCDF. The proposal was recently accepted and will be funded.

- Dependencies, challenges, problems and risks include:
    - Small group (and shrinking) of developers for supporting large project.
    - Dependency on HDF5, controlled by external group.
    - Slow progress in user adoption of netCDF-4 features.

# Ongoing Activities

**We plan to continue the following activities:**

- Provide support to a large worldwide community of netCDF developers and users.
- Continue development, maintenance, and testing of source code for multiple language libraries and generic netCDF utility programs.
- Improve organization of Doxygen-generated documentation for netCDF-C and Fortran libraries.

# New Activities

**Over the next three months, we plan to organize or take part in the following:**

- Prepare material for the upcoming AGU and AMS conferences.
- Release the next versions of netCDF-C, netCDF-Fortran, netCDF-C++.
- Modernize the netCDF documentation to provide easy access to documentation for older versions of netCDF.

**Over the next twelve months, we plan to organize or take part in the following:**

- Begin integration of the upcoming ExaHDF5 features into the netCDF-C, Fortran and C++ interfaces.
- Release an official Windows port of the netCDF-Fortran and netCDF-C++ interfaces.
- Deploy a release with compression competitive with GRIB2.
- Participate in development of new CF 2.0 conventions for climate and forecast simulation output and observational data in netCDF-4 form.
- Continue to encourage and support use of netCDF-4's enhanced data model by third-party developers.
- Create and release online educational material in the form of Youtube video tutorials for using netCDF.

**Beyond a one-year timeframe, we plan to organize or take part in the following:**

- Implement DAP-4 client support in netCDF C library.
- Implement support for Amazon S3 in the netCDF C library.
- Provide thread-safety for the netCDF C library.
- Improve scalability to handle huge datasets and collections.

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

1. **Are there any HDF5 features that you wish netCDF supported?**
2. **Should netCDF be ported to and/or maintained for any other programming computing/development environments?**
3. **How can we encourage more user testing of the release candidates we provide?**

# Relevant Metrics

There are currently about 140,500 lines of code in the netCDF C library source.
The Coverity estimate for defect density (the number of defects per thousand lines of code) in the netCDF C library source has been reduced slightly from **0.35** six months ago to **0.34** today. According to Coverity static analysis of over 250 million lines of open source projects that use their analysis tools, the average defect density with 100,000 to 500,000 lines of code is **0.50**.

Google hits reported when searching for a term such as netCDF-4 don't seem very useful over the long term, as the algorithms for quickly estimating the number of web pages containing a specified term or phrase are proprietary and seem to change frequently. However, this metric may be useful at any particular time for comparing popularity among a set of related terms.

Currently, Google hits, for comparison, are:

- **535,000** for netCDF-3
- **517,000** for netCDF-4
- **384,000** for HDF5
- **82,700** for GRIB2

Google Scholar hits, which supposedly count appearances in peer-reviewed scholarly publications, are:

- **233** for netCDF-3
- **354** for netCDF-4
- **5,920** for HDF5
- **490** for GRIB2

---

Prepared *September 2015*

# Status Report: Python

*April - September 2015*

*Ryan May, Sean Arms, Julien Chastang, Ward Fisher, Russ Rew*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Python can facilitate data-proximate computations and analyses through Jupyter Notebook technology. In particular, Jupyter Notebook web servers can be co-located to the data source for analysis and visualization through web browsers. This capability in turn, reduces the amount of data that must travel across computing networks.
2. **Develop and provide open-source tools for effective use of geoscience data**
   Our current and forthcoming efforts in the Python arena will facilitate analysis of geoscience data. This goal will be achieved by continuing to develop Python APIs tailored to Unidata technologies. Starting with the summer 2013 Unidata training workshop, we developed an API to facilitate data access from a THREDDS data server. This effort has been encapsulated with the new siphon project, which is an API for communicating with a THREDDS server. Moreover, Python technology coupled with HTML5 Jupyter Notebook technology has the potential to address "very large datasets" problems. In particular, a Jupyter Notebook can be theoretically co-located to the data source and accessed via a web browser thereby allowing geoscience professionals to analyze data where the data reside without having to move large amounts of information across networks. This concept fits nicely with the "Unidata in the cloud" vision. Lastly, as a general purpose programming language, Python has the capability to analyze and visualize diverse data in one environment through numerous, well-maintained open-source APIs.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The TDS catalog crawling capabilities found in siphon will facilitate access to data remotely served by the Unidata TDS, as well as other TDS instances around the world. The desired goal of pyCDM is to construct a geoscience focused data model in Python, based heavily on the netCDF-Java implementation of the Common Data Model (CDM). pyCDM is anticipated to provide a simple, pythonic API to the higher level functionality of the FeatureType layer of the CDM.

4. **Build, support, and advocate for the diverse geoscience community**
   Based on interest from the geoscience community, Unidata, as part of its annual training workshop, hosted a three day session to explore “Python with Unidata technology”. Also, to try to help the use of NetCDF in Python, Unidata has promoted Jeff Whittaker’s NetCDF4-python project, including hosting its repository under Unidata’s GitHub account. Unidata is also fostering some community development of meteorology-specific tools under the MetPy grassroots project.

# Activities Since the Last Status Report

## Training Workshop

The Python with Unidata Technologies workshop had 12 attendees, and was again the most well-attended of all the training workshop. This year, we expanded the workshop to 3 days, mostly just to improve pacing of the material; this seemed to work out well. It is interesting to note how many users new to Python were in attendance. It is also interesting to note that attendance was dominated by IT staff and oceanographers; there were not many meteorologists in attendance.

In conjunction with the workshop we also developed a Unidata Python Docker image. It contains a minimal conda distribution along with packages related to Unidata technology and Python.

## Siphon

Siphon represents a rebranding of PyUDL, as we try to elevate our Python support in TDS to a higher status. We anticipate developing Siphon to ensure that it is easy as possible to download data from a TDS in Python, keeping pace with new features added on the Java side.

### Progress has been made on the following:

- Cleaned-up catalog parsing
- Complete implementation of CDM Remote protocol
- Implemented clients for speaking to TDS REST protocols: NCSS and Radar Query Service

## JupyterHub

JupyterHub, part of project Jupyter, is a multi-user Jupyter Notebook server, with a highly-pluggable design. In support of several cloud efforts (NOAA big data, server-side processing), Ryan May has developed a set of docker images that support running a Unidata JupyterHub instance running on Amazon EC2. Authentication of users is managed using GitHub (against a simple whitelist of allowed users). Users are sandboxed from each other (and the master system) through Docker, which allows spawning individual containers on a per-user basis. Facilities provided through the Jupyter Notebook interface include: uploading files (both notebooks and data), terminal access (for installing packages, including using git), and of course execution of Python 2 and 3 code (or potentially other kernels). The interface also works on tablets, giving a nice solution for doing Python analysis through a tablet.

We would like to start extending the testing of this server outside Unidata, to see how this capability solves issues of working with large remote datasets, as well as providing managed Python environments.

**Progress has been made on the following:**

- Implementing GitHub OAuth setup
- Proper SSL certificate for proper functionality on iOS
- Learning how to run Docker containers from within Docker
- Spinning up machines on Amazon EC2 using Docker-machine

**Dependencies, challenges, problems, and risks include:**

- Management of user storage space is tricky, though Docker 1.9 includes actual volume support which should ease this
- This relies upon having the resources to run a sufficiently powerful machine in EC2 as a service to the community

## MetPy

After feedback from the last users' committee meeting, a push was made to bring MetPy forward as a place for community collaboration on meteorology tools that fit within the rest of the scientific Python ecosystem (aka. PyGempak). This project was announced for collaboration in late May with a blog post, and followed up with a presentation at the triennial workshop in June. Feedback has been quite positive, even beyond those who have participated on GitHub.

A presentation for 2016 AMS Python symposium has been submitted, which will hopefully do more to drive event further community interest in this project. Ideas for further development are outlined on GitHub.

**Progress has been made on the following:**

- Project infrastructure created, including automated testing, documentation builds, and conda package generation.
- Skew-T plots, NEXRAD data reading, and unit-aware python calculations all present and well-documented
- Community awareness and involvement progressing well given the early status of the project

**Dependencies, challenges, problems, and risks include:**

- As a grassroots project (without dedicated staff time), it's difficult to consistently keep the project moving forward

## External Participation

The Python team attends conferences as well as participates in other projects within the scientific Python ecosystem. This allows us to stay informed and to be able to advocate for our community, as well as keep our community updated on developments. Ryan May attended the 2015 SciPy conference in Austin; major takeaways:

- xray is a new package from Stephen Hoyer of the Climate Corporation; it serves as a layer on top of netcdf data to provide simple query capabilities (think pandas for multidimensional arrays) There was a lot of excitement from the SciPy community regarding this package
- Project Jupyter announced another major round of funding, which ensure that the Jupyter notebook will continue its steady development
- There was a distinct lack of meteorologists at the conference, in comparison with the size of the oceanography community present. We would encourage anyone with interest in Python to consider attending the conference; it's a fun and really informative week.

Ryan May has also continued to be an active participant in the matplotlib community, reviewing some pull requests and contributing several others. We also continue to host Jeff Whittaker's netCDF4-python project repository; Jeff continues to be the active maintainer of the project.

**Progress has been made on the following:**

- Fixed unit support in matplotlib--this facilitates improved unit support for MetPy's Skew-T plotting
- Contributed support in matplotlib for embedding animations (as HTML5 video) within the Jupyter Notebook
- Contributed a pull request to merge in the community developed JSAnimation package, which embeds animations in the notebook as a javascript animation.

**Dependencies, challenges, problems, and risks include:**

- Due to little dedicated staff time for these activities, keeping up on these activities is not guaranteed

# Ongoing Activities

**We plan to continue the following activities:**

- "Python with Unidata Technologies" training workshop
- Maintaining Siphon as an official Python API for working with TDS
- Continued participation in the scientific Python community
- Relevant matplotlib support and fixes
- Working with JupyterHub as a way to facilitate data-proximate analysis
- Growing and developing MetPy as a community resource for Python in meteorology

# New Activities

**Over the next three months, we plan to organize or take part in the following:**

- Using supplemental funds from NSF, develop asynchronous training materials for Python in meteorology. We are investigating the use of a cloud server hosting executable Jupyter Notebooks (based on our training workshop) as the core of the

training materials, using either the tmpnb or jupyterhub packages from Project Jupyter.

**Beyond a one-year timeframe, we plan to organize or take part in the following:**

- Evaluate the possibility of extending siphon functionality to interface with the AWIPS-II EDEX server

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

1. What are the biggest obstacles that you see to the use of Python with other Unidata technologies, or for use in meteorology in general?
2. How valuable do find an effort like MetPy to the Python meteorology community? Are there additional barriers we could remove through this project? Are there other efforts over which this should take priority?

# Relevant Metrics

Siphon (since April):

- 94% test coverage
- 544 downloads/month from the Python package index (no metrics for anaconda.org)
- 3 externally contributed issues, 1 external pull request

MetPy (since April):

- 95% test coverage
- 847 downloads/month from the Python package index (no metrics for anaconda.org)
- 5 externally contributed issues, 3 external pull requests

---

Prepared *September 2015*

# Status Report: Support

*April - September 2015*

*Tom Yoksas, Jennifer Oxelson, UPC Staff*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *Unidata User Support enables access to geoscience data by supporting the use of tools created and/or supported by the UPC.*

2. **Build, support, and advocate for the diverse geoscience community**
   *The user support provided by the UPC is recognized throughout the atmospheric science community. Unidata's outreach efforts are routinely called out in surveys of the NCAR/UCAR community.*

## Activities Since the Last Status Report

### Training

- The UPC and the University of South Florida (USF) hosted a regional training workshop focusing on use of the IDV, RAMADDA and AWIPS-II at the USF on April 10-12, 2015. Unidata staff participating in the training were Michael James (AWIPS-II), Yuan Ho (IDV), and Tom Yoksas (IDV, RAMADDA, LDM).
- The UPC hosted the Unidata Users triennial workshop entitled **Data-Driven Geoscience: Applications, Opportunities, Trends, and Challenges** at the UCAR Center Green campus in Boulder, CO on June 22-25, 2015.
- The UPC hosted its annual training workshop series from July 20 - August 5, 2015 in its Foothills Lab offices in Boulder, CO.
- The UPC hosted a regional training workshop focusing on the use of the IDV, McIDAS-V and RAMADDA by Central American Weather Service forecasters in San Jose, Costa Rica. Unidata staff participating in the training were Yuan Ho (IDV and McIDAS-V) and Tom Yoksas (McIDAS-V and RAMADDA). This activity was funded by USAID.

## New Activities

**In order to fulfill our objectives articulated in the Unidata 2018 Proposal, focused efforts are needed in two major areas:**

- Enhance electronic support offerings
- Create instructional materials for online virtual training

# Relevant Metrics

Since January 26, 2006 over 46660 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system.

**Support by Category**

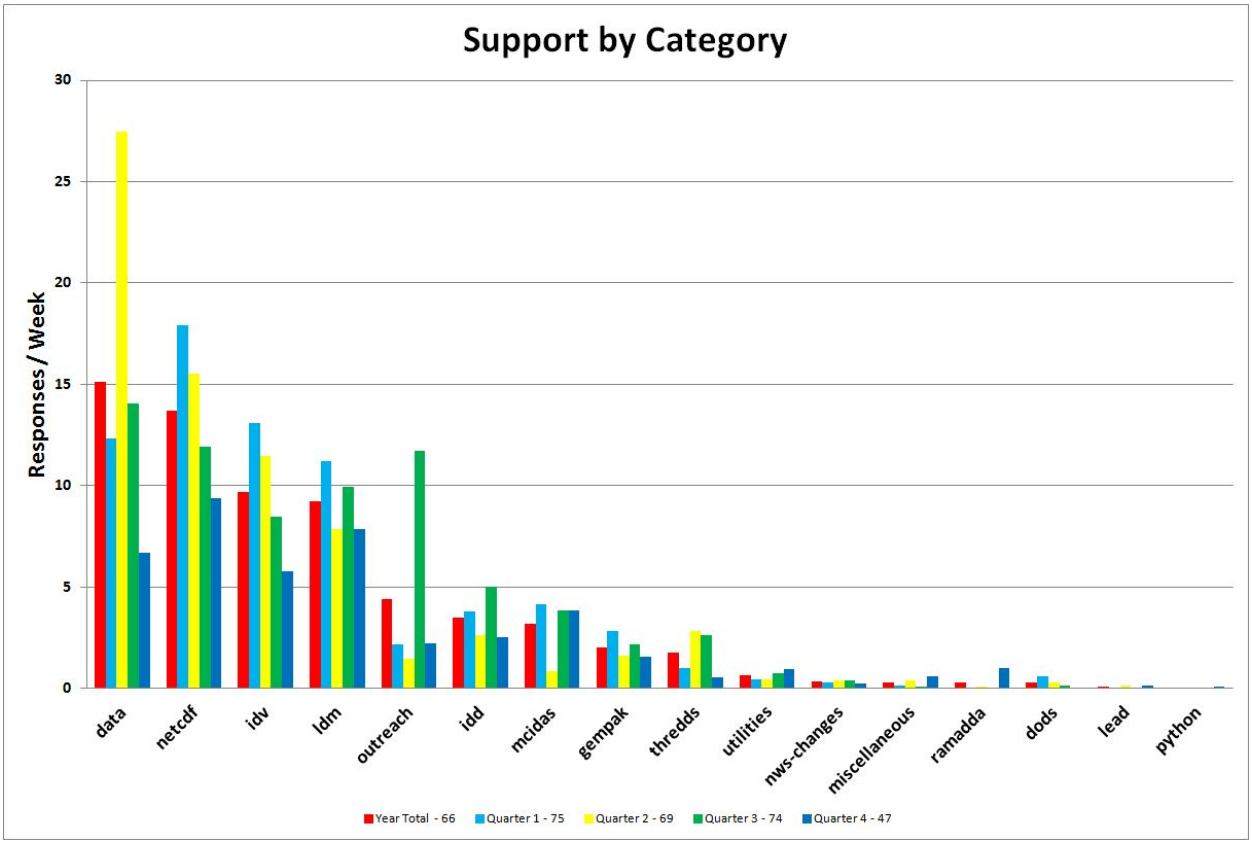


**Fig. 1:** Above are histograms that portray the number of Unidata email responses for categories of support for a one year period ending September 15, 2015. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period is depicted from oldest to newest from left to right. The number of responses has been normalized to weekly averages so that the support load over the various periods can be easily compared.

Individual support activities included in the categories depicted above are listed in the following table.

| Category | Packages, Groups, and Lists |
|---|---|
| **data** | casestudies, casestudies-list, conduit, craft, craft-ty, craft-nws, datastream, difax, eumetsat, level2, level2-ty, noaaport, |

| | |
|---|---|
| | noaaport-ty, noaaportldm |
| **dods** | dods, dods-core, dods-list, dods-tech, dods-mlgui-tc, dods-pm, dods-tac, dods-team, opendap, opendap-core, opendap.forward, opendap-list, opendap-tech |
| **gempak** | gempak, gembud-list, gempak-ty, awips-ty |
| **idd** | idv, idvlist, idvsteering, java-gui, metapps, visad, visad-list, visad-renderer |
| **idv** | idv, idvlist, idvsteering, java-gui, metapps, visad, visad-list, visad-renderer |
| **ldm** | ldm, ldm-users-list |
| **lead** | lead, leadusers |
| **mcidas** | mcdevelop, mcdevelop-ty, mcidas, mcidas-list, mcidas-ty |
| **miscellaneous** | esupport, fxlinux, misc, license, network, notrack, platforms, wxp wxp-lis |
| **netcdf** | data-models, libcf, ncml, netcdf, netcdf-miss, netcdfgroup-list, netcdf-hdf, netcdf-hdf-list, netcdf-java, netcdf-perl |
| **nws-changes** | nws-changes |
| **outreach** | agu-ty, announce, argentina-ty, barbados-ty, brazil-ty, cathalac-ty, chile-ty, costarica-ty, mexico-ty, support-ty, cbmet-ty, community-list, eletter, egrants, eumetsat-ty, external, iai-ty, international-ty, joss-ty, k12-list, korea-ty, meteoforum-ty, unidata, workshop |
| **python** | python |
| **ramadda** | ramadda |
| **thredds** | java-dev, java-dev-list, thredds |
| **utilities** | decoders, ldm-mcidas, udunits |

**Support by Topic**

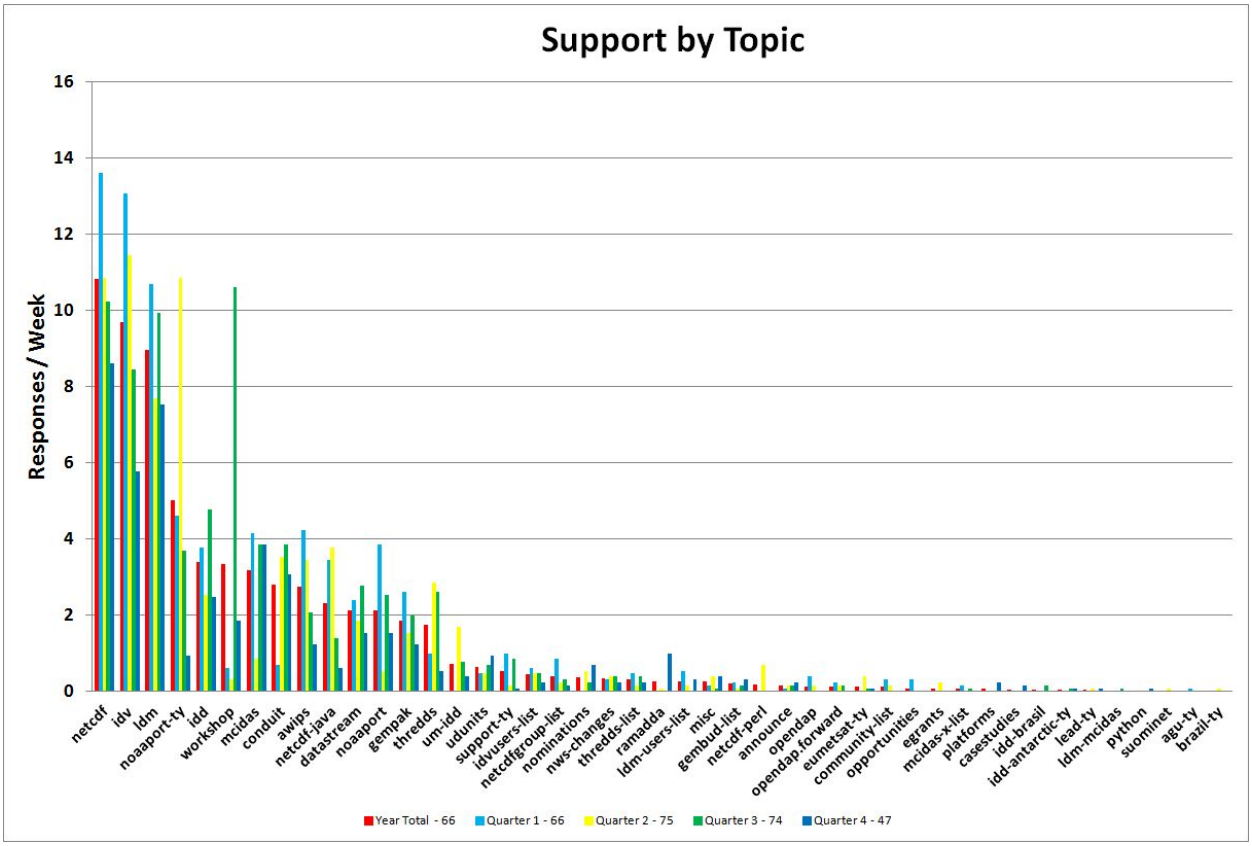


**Fig. 2:** Above are histograms that portray the number of Unidata email responses for individual topics of support for a one year period ending September 15, 2015. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period is depicted from oldest to newest from left to right. The number of responses has been normalized to weekly averages so that the support load over the various periods can be easily compared.

# Comments

- The total support provided by the UPC remains high, and yearly totals have been relatively constant for the past two years. Overall support activities vary by somewhat by quarter. Spikes in support for individual packages is largely correlated with the release of new distributions, and, for the IDV in particular, jumps after training workshops.
- Support for netCDF continues to be substantial, and is understandable given the **large** number of users of the package worldwide.
- The IDV support load is second only to that for netCDF; no large increases have been seen over the past 6 months.
- Support for netcdf-java continues to grow steadily.
- Support for the legacy visualization packages GEMPAK and McIDAS continues to be

substantial.  Support for AWIPS-II has been increasing steadily and now exceeds that for GEMPAK.
- Support for LDM, IDD, and data continues at a high level and shows some variability throughout the year.
- Taken as a whole, the support required for visualization packages (GEMPAK, IDV, McIDAS and AWIPS-II) is comparable to the support related to data ingest and distribution (LDM, IDD, noaaport).
- The high numbers for outreach reflect the high level of activity in a variety of activities including organizing sessions at various national meetings.

# Notes

These numbers and conclusions should not be taken too literally, for several reasons:

- For some packages, multiple responses in the same thread may be bundled into a single archived email. Other packages have each response in a thread counted separately.
- After a new release of software, there may be a flurry of the same or similar questions, which can be answered in separate emails or in a single mailing list posting.
- The graph primarily represents support of end users and site administrators, not developers. Support for non-Unidata developers in projects such as THREDDS, IDV, GEMPAK, and McIDAS requires significant resources, but is difficult to assess.
- Not all support records were indexable for this report. Given this, the above numbers are an ***underestimate*** of the actual support being provided by the UPC.

Additional User Support Inquiry Metrics

---

Prepared *September, 2015*

# Status Report: THREDDS

*April - September 2015*

*John Caron, Sean Arms, Ethan Davis, Dennis Heimbigner, Ryan May, Christian Ward-Garrison*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   The work of the THREDDS group is comprised of two main areas: the THREDDS Data Server (TDS) and the Common Data Model (CDM) / netCDF-Java library. The TDS provides catalog and data access services for scientific data using OPeNDAP, OGC WCS and WMS, HTTP, and other remote data access protocols. The CDM provides data access through the netCDF-Java API to a variety of data formats (e.g., netCDF, HDF, GRIB). Layered above the basic data access, the CDM uses the metadata contained in datasets to provide a higher-level interface to geoscience specific features of datasets, in particular, providing geolocation and data subsetting in coordinate space. The CDM also provides the foundations for all the services made available through the TDS.

   The data available from the IDD is a driving force on both the TDS and netCDF-Java development. The ability to read all the IDD data through the netCDF-Java library allows the TDS to serve that data and provide services on/for that data.

2. **Develop and provide open-source tools for effective use of geoscience data**
   Unidata's Integrated Data Viewer (IDV) depends on the netCDF-java library for access to local data, and on the THREDDS Data Server (TDS) for remote access to IDD data. At the same time, the CDM depends on the IDV to validate and test CDM software. Many other tools build on the CDM / netCDF-Java library (eg ERDDAP, Panoply, VERDI, etc) and on the TDS (ESGF, LAS, ncWMS, MyOcean, etc).

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The Common Data Model (CDM) / netCDF-Java library is one of the few general-purpose implementations of the CF (Climate and Forecast) metadata standards. Current active efforts in CF that we are involved with include use of the extended netCDF-4 data model (CF 2.0) and for point data (Discrete Sampling Geometry CF-DSG).

   The TDS has pioneered the integration of Open Geospatial Consortium (OGC) protocols into the earth science communities. Strong international collaborations have resulted in WCS and WMS services as part of the TDS.

   The CDM and TDS are widely used implementations of the OPeNDAP DAP2 data access protocol. Unidata has worked with the OPeNDAP group to design, develop, and implement a new version of the DAP specification, DAP4, which is now available

in the TDS server and the netCDF-Java client software stack.

4. **Build, support, and advocate for the diverse geoscience community**
   The THREDDS project is involved in several international standardization efforts (CF, OGC, etc.) which cross-cut a multitude of disciplines, both inside and outside of the geoscience community. The netCDF-Java client library, as well as the TDS often serve as incubators for new pushes in these efforts.

# Activities Since the Last Status Report

## The THREDDS Project

The THREDDS Project encompases four projects: netCDF-Java, the THREDDS Data Server (TDS), Rosetta, and siphon (the Unidata Python client to interact with a TDS server).

## Released netCDF-Java / TDS version 4.6 (Stable)

The stable release of both netCDF-Java and the THREDDS Data Server is version 4.6.

### Progress has been made on the following:

- GRIB Collections now scale to large numbers of files.
- We now use the Gradle build system
- Using Coverity to find and fix more than 4000 defects. Defect count now < 1 / 1000 Lines of Code.

### Dependencies, challenges, problems, and risks include:

- Addressing feedback as community upgrades their TDS installations

## 2015 TDS Training Workshop

The 2015 TDS Training Workshop utilized Docker container technology which freed up time for teaching more TDS-specific material.

### Progress has been made on the following:

- Documentation updated to reflect latest changes in version 4.6.
- Used Docker for running the TDS in the workshop

### Dependencies, challenges, problems, and risks include:

- Major changes coming in 5.0 - will require another very thorough pass over documentation, training materials, etc.

# Ongoing Activities

**We plan to continue the following activities:**

- Documentation updates - reworking the tutorial material to use Docker
- Maintain thredds.ucar.edu and keep up with the addition of new datasets to the IDD
- Continue development of the TDS python client siphon, as well as potentially extend its functionality to interface with the AWIPS-II EDEX server
- Continue to implement a Rosetta interface for each discrete sampling geometry (DSG) from the CF-1.6 specification (http://cfconventions.org/Data/cf-conventions/cf-conventions-1.6/build/cf-conventions.html#discrete-sampling-geometries)

**The following active proposals directly involve THREDDS work:**

- New EarthCube award: "Advancing netCDF-CF for the Geosciences". This two year, Unidata lead project will work to extend netCDF-CF conventions in ways that will broaden the range of earth science domains whose data can be represented.
- Beginning the second year of NASA ROSES ACCESS award: "High Performance Multidisciplinary Open Standard Data Services to Serve Terrestrial Environmental Modeling" with USGS CIDA.
- Three EarthCube awards are finishing up on a no-cost extension:
  1) EarthCube Building Blocks award: "Integrating Discrete and Continuous Data" with Univ of Texas, Austin and others.
  2) EarthCube Building Blocks award: "Specifying and Implementing ODSIP, A Data-Service Invocation Protocol" with OPeNDAP, Inc.
  3) EarthCube Building Blocks award: "Deploying Web Services Across Multiple Science Domains" with IRIS, UNAVCO, and others. Period of performances: Oct 2013 - Sept 2015.
- Two NASA ROSES ACCESS proposals were submitted this year:
  1) "Interactive Algorithm Development and Product Validation through Innovative Data Access and Visualization Methods" with UWisc/SSEC.
  2) "Leveraging available Technologies for Improved interoperability and visualization of Remote Sensing and in-situ Oceanographic Data at the PO.DAAC" with JPL/PO.DAAC

# New Activities

**Over the next three months, we plan to organize or take part in the following:**

- Getting netCDF-Java/TDS v5.0 to a beta release
- Making public a TDS 5.0 Test Server
- Enable TDS to serve data from Amazon S3 buckets
- Finalize visualization preview of converted data in Rosetta
- Move issue tracking, roadmap planning, etc. from our Jira server to GitHub
- Host non-Maven artifact downloads (i.e. toolsUI.jar, netCDFAll.jar, tdm.jar, ncIDV.jar, and thredds.war) on github.

**Over the next twelve months, we plan to organize or take part in the following:**

- Update the TDS 5.0 ncWMS plug-in to use the new ncWMS 2.0 code from the University of Reading
- Upgrade the ncWMS, ncISO, and other plugin services to use the new TDS 5.0 plugin layer
- Transitioning thredds.ucar.edu to TDS 5.0
- Getting netcdf-Java/TDS v5.0 to a stable release
- Incorporate ncSoS into THREDDS 5.0
- Enable siphon to interface with CDMRFeature objects

**Beyond a one-year timeframe, we plan to organize or take part in the following:**

- Integrate Rosetta with the TDS to allow conversion and publishing of observational ASCII-based datasets to the TDS
- Move to a fully online based training tutorial, reserving in-person, annual training for advanced topics

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

1. **Does your department or campus IT utilize Docker technology?**

2. **Have there been discussions regarding moving student computing resources (i.e. computer labs) to the cloud?**

3. **What are the top three analysis and visualization tools utilized in a) the classroom, b) student research, and c) faculty research**

# Relevant Metrics

While it is still early in the semester, it should be noted that the top client accessing data from the THREDDS data server over the past month is now Python. The THREDDS team will keep a close eye on our server usage statistics to see if this continues to be the case, and will provide a more detailed report for the Spring meeting.

---

Prepared *September 2015*

# Status Report: Outreach Activities

*April - September 2015*

*Ben Domenico*

## Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *Work with representatives of other disciplines and serve on their governing boards where appropriate, e.g., NCAR GIS for Geographic Information Systems, CUAHSI (Consortium of Universities for Advancement of Hydrological Science), ODIP (Ocean Data Interoperability Platform), etc.*
2. **Develop and provide open-source tools for effective use of geoscience data**
   *Work with Unidata collection of Ipython notebooks in cloud development platforms and on native Microsoft Windows.*
3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   *Continue to work with Opengeospatial Consortium (OGC) to augment international CF-netCDF standards that have been established over the last several years.*
4. **Build, support, and advocate for the diverse geoscience community**
   *Serve as Co-Investigator on Earthcube Cyberconnector project which will make Unidata data available to a wide range of research and education communities beyond the traditional Unidata community.*

## Activities Since the Last Status Report

### Open Geospatial Consortium activities

*Continue to work with Opengeospatial Consortium (OGC) to augment international CF-netCDF standards that have been established over the last several years.*

#### Progress has been made on the following:

- Represented Unidata at OGC Technical Committee meeting
- Chaired OGC NetCDF Standards Working Group
- Served as UCAR business representative to the OGC.
- Represented Unidata at ODIP Steering Committee meetings and telcons
- Hosted the OGC Technical Committee meetings at UCAR/Unidata in June.

#### Dependencies, challenges, problems, and risks include:

- Tight travel fund situation is reducing participation in OGC Technical Committee meetings, e.g., the TC meeting this month. The NetCDF SWG was ably chaired by

Lorenzo Bigagli of ESSI Labs at the University of Florence.
- The ODIP 2 proposal was funded by the European Commission, but not by the US NSF. Here also, lack of funds prevents ongoing active participation in ODIP workshops.
- Windows 7 Ipython notebook failure needs to be resolved.

## EarthCube Cyberconnector Project

Collaborative project with George Mason University to make Unidata real time datasets available to researchers and educators in other disciplines

### Progress has been made on the following:

- Introduced GMU technical staff to Unidata technical staff in during OGC Technical Committee meetings.
- Held kick off teleconference to outline steps for getting a "motherlode clone" running on the GMU EarthCube hardware.

### Dependencies, challenges, problems, and risks include:

- The project got off to a slow start but the PI, Liping Di of GMU, anticipates that a one year no cost extension will make it possible to achieve the proposed objectives.
- During the kick off telecon, some issues relating to the special port used by the Unidata LDM cropped up but GMU technical staff are pursuing solutions.

## Python Workshop Notebooks on alternative platforms

*Work with representatives of other disciplines and serve on their governing boards where appropriate, e.g., NCAR GIS for Geographic Information Systems, CUAHSI (Consortium of Universities for Advancement of Hydrological Science), ODIP (Ocean Data Interoperability Platform), etc.…*

### Progress has been made on the following:

- Attempted to get Python Workshop Ipython notebooks running in the cloud development environment hosted by Wakari. This failed because Wakari does not yet support Ipython 3.x.
- Got workshop Ipython notebooks running on Windows 10. Most of them worked fine. There were some problems with operating system commands that are not Windows native and with the way Windows handles temporary file permissions.
- Could not get workshop Ipython notebooks running at all on Windows 7 machine

### Dependencies, challenges, problems, and risks include:

- Need to revisit workshop notebooks on Wakari when they support Ipython 3.x which

is required for this years workshop notebooks
- If time allows, it will be good to clean up the glitches in the notebooks on native Windows 10.
- Windows 7 Ipython notebook failure needs to be resolved.

# Ongoing Activities

**We plan to continue the following activities:**

- **Coordination and collaboration with NCAR GIS**
- **Represent Unidata and UCAR in OGC**
- **Participate in ODIP 2 to the extent possible**
- **Participate in Earthcube Cyberconnector project.**

# New Activities

**No new activities are planned. For the next year, the objective is to keep current activities going as well as possible and make arrangements to complete those that can be completed. Note that, in response to the budget crunch, Ben is reducing his Unidata FTE commitment.**

# Areas for Committee Feedback

**We are requesting your feedback on the following topics:**

> **How important is it to ensure that Ipython notebook development can be done in the cloud and on Windows platforms?.**

---

Prepared *September 2015*