

Unidata Users Committee Meeting

September 26-27, 2016
Unidata Program Center

Agenda

Monday, 26 September 2016

8:30-9:00 Continental Breakfast

9:00-9:15 Administrative Items

- Introduction: New Members
- Date for Spring meeting
- Review of [Action Items](#) from April 2016 meeting

9:15-9:30 Strategic Advisory Committee Report - Mohan Ramamurthy and Russ Schumacher

9:30-10:45 Director's Report & Questions - Mohan Ramamurthy

10:45-11:00 Break

11:00-11:45 Around-the-table reports from members

11:45-12:45 Lunch

12:45-1:15 Status Reports - (Users Committee should review status reports prior to the meeting) - staff will be available for questions

1:15-2:00 Python updates - Ryan May

2:00-2:15 Break

2:15-3:15 NOAA UPDATES: CONDUIT & AWIPS - Becky Cosgrove

3:15-4:00 Update on Docker Activities - Julien Chastang

4:00-4:15 Transition to FL2 Auditorium

4:15-5:00 2014 DeSouza Award Lecture

5:00 Adjourn

6:30 Collaborative discussion on the day's proceedings over dinner

- Dinner at the Rio Grande ([Map](#))

Tuesday, 27 September 2016

8:39-9:00 Continental Breakfast

9:00-9:25 Strategic Plan Update - Josh Young/Mohan Ramamurthy

9:25-9:50 Big Data Project Update - Jeff Weber

9:50-10:15 Online Python Tutorial/DMRC Update- Josh Young/Larissa Gordon

10:15-10:30 Break

10:30 - 11:00 AWIPS Update - Michael James

11:00 - 11:45 Blue Sky Session

11:45-12:00 Wrap-up session, review/assign Action Items

12:00 Adjourn

Status Report: AWIPS and GEMPAK

April - September 2016

Michael James

AWIPS Activities Since the Last Status Report

AWIPS in the Cloud

The \$20k Azure for Research Grant for Unidata AWIPS has allowed for the edex-cloud open data server to live on, with an on-site EDEX server available as a replacement for those periods of time where a cloud-based server is not funded.

This grant has also enabled the development of a RedHat 7 supported EDEX and CAVE build, which can take advantage of the Azure file sharing architecture to create a distributed EDEX environment, scalable to data requirements.

Notes about cloud-based AWIPS:

- Unidata has staged a development and testing EDEX server in the Azure cloud running on Red Hat/CentOS 7.
- The edex-cloud data archive was expanded to retain 30 days of NEXRAD3 and FNEXRAD products.
- A talk was submitted to the 97th AMS Annual Meeting on the development and deployment of a distributed EDEX server in the Azure cloud.
- EDEX 16.2.2 can install and leverage GEMPAK/nex2gini for real-time composite creation in the cloud (supplanting and expanding the FNEXRAD product set currently generated in AWS).
- Published [Jupyter notebooks](#) which provide access to real-time AWIPS data in the cloud, using the [python-awips](#) API.

Data Delivery / OGC Plugin Investigation

Unidata has recently received access to the awips2-data-delivery and awips2-ogc repositories hosted on NCEP's Vlab. Data Delivery gives users the ability to create queries and subscriptions to data sets hosted on OpenDAP servers such as NOMADS and THREDDS. For NWS operations, the Data Delivery Central Registry is run at NCEP, which is then subscribed to by downstream WFO offices, with subscriptions available to the CAVE client. A similar model is being investigated for Unidata AWIPS, to enable clients to subscribe to any OpenDAP data catalog.

OS X Support

Unidata has supported CAVE for OS X since 15.1.1, though support for all viz plugins and perspectives is still in progress. In 16.2.2, the NSHARP and GEMPAK libraries have been completely rebuilt from GEMPAK FORTRAN and C source files, and viz.core.gl plugins have been further refactored to prevent erratic behavior when disposing/destroying GL contexts. The biggest remaining challenge is to refactor XML serialization / marshalling in NCEPs' National Centers Perspective, which presently is unsupported and prevents the entire NCP from loading.

The next major step in OS X AWIPS is to package and sign the Cave application for distribution in Apple's App Store through UCAR's Apple developer account.

Windows Support

Taking advantage of Microsoft's free MSEdge Windows 10 virtual machines, a new Windows client for 16.2.2 is in development. For this release I requested and received a open-source project license for Advanced Installer (<http://www.advancedinstaller.com>).

Python Data Access Framework (python-awips)

In March 2016 I published a [standalone Python Data Access Framework](#) (DAF) for retrieving data from an AWIPS EDEX server. This package allows users to create a data request to be sent to a remote EDEX data server (examples use edex-cloud.unidata.ucar.edu). The interface will return Python data objects which are easily convertible to Numpy arrays, to be rendered with existing Python packages such as Matplotlib/Basemap, Cartopy, and MetPy.

Unidata AWIPS Release Notes

16.2.2 TBD (Sep-Oct 2016)

- The LDM can now be started and stopped by both users 'awips' and 'root', solving the process ownership conflict which happened when root issued an "ldmadmin start|stop" command after the service was already started by 'awips'.
- NAWIPS-like colorized surface temperature bundle added to the Surface Obs menu.
- Using the bundled GEMPAK binary release, the awips user cron will now create FNEXRAD composites of DAA (one hour precip) and DTA (storm-total precip), which are then immediately ingested. For now this workflow provides national composites which are unavailable on the FNEXRAD feed coming out of AWS.
- LDM scour is no longer used to purge raw data files, replace by the EDEX archive plugin purge process run at 30 minutes past the hour, every hour. LDM would scour down to 24 hours, while EDEX scours everything to 1 hr, freeing 100 GB of disk

space.

- Updated NUCAPS/NPP NSHARP plugin which now colors sounding locations with quality control flag.
- New NEXRAD Display plugin (RadarMapResource) for D2D with selectable stations to load level 3 site products and map bundles.
- NCEP's Day-Night Terminator overlay resource now available to plot in D2D perspective.
- Re-tooled the creation of local maps for every WFO and 88D site by leveraging CAVE's projection tool to calculate and export geometries and write bundle XML files for each station.
- Added a restart button to the CAVE Localization Preferences page, made active to the user when site or localization info changes.
- Mac OS X client features re-build of NSHARP library for skew-t display (Raob, NPP, forecast models).
- New RPM builds for python (2.7.10) and numpy (1.9.2)

16.1.5 June 2, 2016

- Switching the localization site in CAVE Preferences now updates the side-view maps to the new WFO area of responsibility.
- New general "awips2-localization" RPM obsoletes site-specific "awips2-localization-OAX" with support for all WFOs, programmatically creating map scales for each WFO office.
- Full support for UNIWISC McIDAS AREA files, including new support for
 - 20km Global Rectilinear IR and WV
 - 30km Global Mollweide IR and WV
- Fix for GINI projection error in decoding of Supernational composites
- New default map projections include Global Mercator, Global Mollweide, and Regional displays for Alaska, Africa, Antarctica, Aus/NZ, Europe, Hawaii, Japan, Pacific Ocean, Puerto Rico, and South America.
- New package "awips2-python-gfe" makes available various Python functions used by GFE which were previously absent.
- Hold-right-click menu on product stack now includes Interpolation toggle.
- Default display of FNEXRAD composites using data interpolation and decreased tiling for clearer data display.
- Two-pane toggle added to D2D right-click menu (along with existing 4-pane toggle)
- Two-pane displays are now oriented side-by-side rather than above-and-below.
- Reconfigured D2D NSHARP plugin to display selectable stations
- Reconfigured image export with default timestamp in file name to prevent overwriting. Files are now exported to /awips2/export/<USER>/
- GFS20 replacing GFS40 in the Models menu
- New METAR precip total plugin colorizes rainfall amounts.
- Reorganization of Surface WATCH/WARN bundles, loadable as TStorm/Tornado, Flood, Marine, Extreme Wines, and "all". Bundle polygon contrast has been increased as well.

- Redbook graphics bundles displayed with higher contour contrast
- Complete regional imagery Satellite sub-menu with full product set for Alaska, Hawaii, Puerto Rico, and Arctic composites.
- New Satellite sub-menus for NOAAport GINI products (GOES East, GOES West, GOES E/W 10 km Composite, 24 km National Composite, 24 km NH Composite, Sounder Imagery)
- NPP VIIRS Near Constant Contrast bundle composites all available area for product display.
- ESRL/GSD experimental 3 km HRRR now supported (HRRRX), ingesting on edex-cloud.unidata.ucar.edu (though still missing some grib parameters)
- Better httpd logging removes the unnecessary reporting of internal connection (responsible for approx. 90% of httpd log file size), and logs hostname of client machine, if available.
- Default purge rule for satellite products preserve 96 frames (12 day archive for Global IR/WV composites), and 288 frames (24 hours) for FNEXRAD images.
- GFE server support available for all WFOs, not just OAX (requires a change to setup.env and edex_camel restart - this does not mean support for all server sites simultaneously)
- Binlightning, nctaf, and nctext decoders re-enabled.
- Removed all metar/synop-to-Shef processing due to performance problems.
- CAVE operates strictly in practice mode, preventing WarnGen products from being pushed to a connected EDEX server.
- GEMPAK can now be installed from the AWIPS yum repository (simply "yum install gempak").

Unidata GEMPAK 7.3.0 Release (May 2016)

- Released as source code tarball and 64-bit Linux RPM May 2016.
- Incorporated NAWIPS 7.3.0 table, map, and code updates.
- Added GFS20 (GFS215) to tables and LDM file pqact.gempak_decoders_grid
- Added a fix for dcwarn, dcwatch, and other decoders which would write empty files when built with gfortran 4.6+
- G2CONV added as a parameter to gdgrib2 for bulk processing GEMPAK parameters.
- NAGRIB2 analysis grid fix for GDATTIM, defined as YYMMDD/HHHH (fixes problem where if grib1 was converted to grib2 and processed with nagrib2, GDATTIM would be incorrectly set as YYMMDD/HHHHF000).
- Fixed negative grid resolution when gdgrib2 converts a GEMPAK grid to grib2. Caused either when 1) `r lon1 > 0 and r lon2 = 0.0`, i.e., eastern hemisphere, and 2) `r lon1 < 0. && r lon2 < 0. && r lon1 > r lon2 .`

Questions for Committee Members

- D2D or NCP?
- What teaching and research needs can the three other perspectives serve if given Unidata's attention? The three being: GFE (Graphical Forecast Editor), Hydro Perspective, MPE (Multisensor Precipitation Estimator) Perspective.
- What products would you like to see in AWIPS that are not already available? Himawari? Global VIIRS? Landsat? GIS datasets? Mesonets? (And how can Unidata get these data?)
- Are faculty and staff at your institution aware that AWIPS can be used in the classroom without installing and maintaining your own EDEX server?
- Are faculty and staff at your institution aware of the Mac and Windows AWIPS clients?
- Should Unidata establish a long-term AWIPS data archive? at and in cooperation with a member institution?
- What shapefiles and maps could be added to AWIPS?
- Any interest in an AWIPS regional workshop?
- Is there something AWIPS does not do that it should?

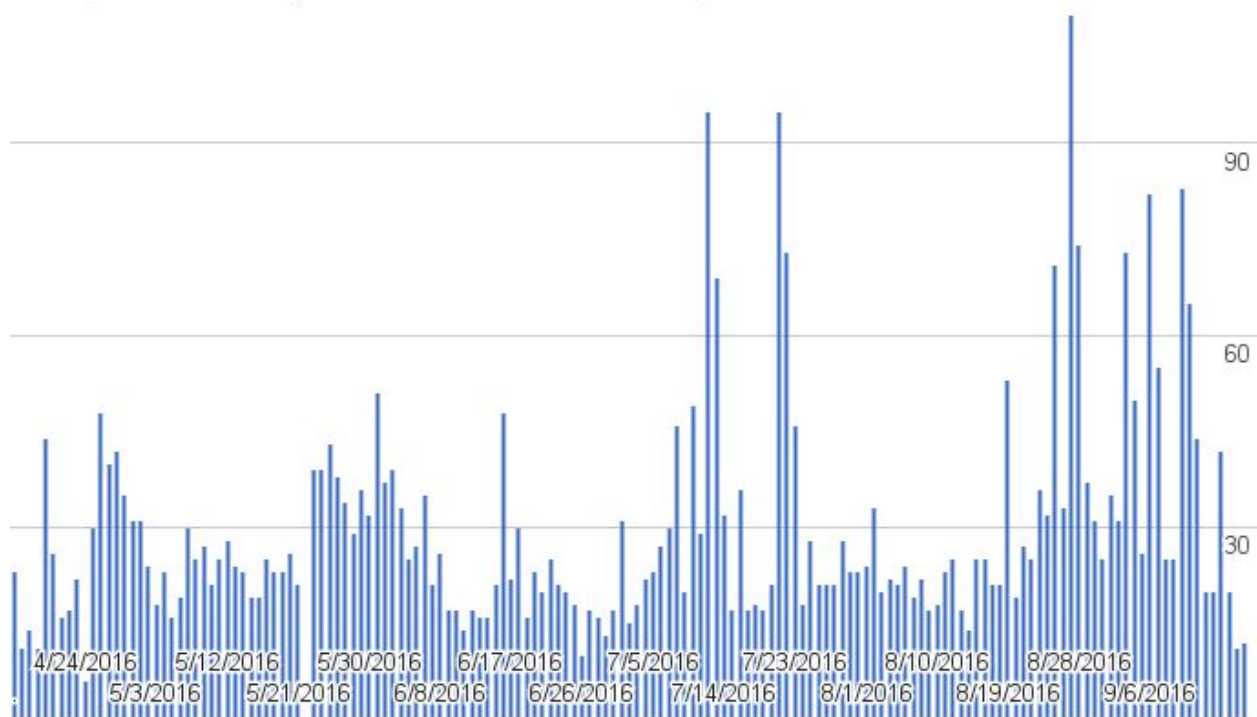
Relevant Metrics

Software Downloads

- [AWIPS Downloads by Month, Jan-Aug 2016](#)
- [GEMPAK Downloads by Month, Jan-Aug 2016](#)
- [AWIPS Downloads by Affiliation](#)
- [AWIPS Downloads US vs. non-US](#)
- [AWIPS Downloads Core vs. non-Core](#)
- [AWIPS Downloads by site for Core Institutions, Jan-Aug 2016](#)
- [AWIPS Downloads by site for EDUs, Jan-Aug 2016](#)
- [more...](#)

Daily AWIPS users (edex-cloud.unidata.ucar.edu)

120



Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

- 1. Enable widespread, efficient access to geoscience data**
Both AWIPS and GEMPAK are freely available, and both incorporate LDM/IDD technology for efficiently accessing geoscience data. The cloud-based data server maintained by Unidata invites a number of users and we are seeing more adoption with time, as the software becomes more stable, more platforms are supported, more data is served conveniently and freely to the community.
- 2. Develop and provide open-source tools for effective use of geoscience data**
Both AWIPS and GEMPAK are open-source, and while GEMPAK is now in maintenance mode, AWIPS is continuously being developed.
- 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 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.

Prepared *September 2016*

Status Report: Cloud Computing Activities

April 2016 - September 2016

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

Unidata technical staff have deployed experimental and production software in several cloud computing environments. For the past two and a half years, Unidata-created products for the IDD, FNXRAD and UNIWISC data streams have been created in a VM hosted in the Amazon cloud. In collaboration with Unidata, [NOAA is delivering 20+ years of NEXRAD Level II data](#) via Amazon Web Services. LDM and [THREDDS Data Server](#) (TDS) software are being employed to deliver these data. In addition, we have an experimental “motherlode” class server running in the Azure cloud serving a subset of the IDD data via a [TDS](#) and [RAMADDA](#). These data are supplied by an LDM relay also running on the Azure cloud.

Activities Since the Last Status Report

Docker Development

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, LDM, ADDE, RAMADDA, THREDDS and Python with Unidata Technologies, 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 various Docker-related Unidata [github](#) and [dockerhub](#) repositories and cloud demonstration servers.

Progress has been made on the following:

- **LDM:** Developed an [LDM container](#) that is currently running in the Azure cloud as a relay node and leaf node and supplying data to [RAMADDA](#) and [THREDDS Data Server](#) containers.
- **RAMADDA:** Collaborated with Unidata Users Committee member Warren Pettee to enhance a [RAMADDA Docker container](#). [This container is running without issue on the Azure Cloud.](#)

- **THREDDS Data Server and TDM:** Currently collaborating with Axiom Data Science to construct a [TDS/TDM container](#) that will satisfy Unidata and Axiom objectives and avoid duplication of effort. We have a [demonstration server running on the Azure Cloud](#) that will soon be replaced by the efforts of this collaboration.
- **ADDE:** We had a demonstration server running for AMS 2016. We have informally asked UW/SSEC to allow us to make freely available a containerized distribution of the ADDE server component of McIDAS-X. A formal petitioning is anticipated around the time of this fall's Strategic Advisory Committee meeting.
- **CloudIDV:** We are in the process of obtaining feedback from beta users.

Dependencies, challenges, problems, and risks include:

While these efforts are promising initial steps, there are challenges ahead in making these technologies useful to our community. Apart from client technologies like the CloudIDV and Jupyter notebooks, 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. Moreover, we may have to rethink workflows in a cloud environment (data-proximate analysis and visualization, for example) in addition to porting present Unidata cyberinfrastructure to the cloud.

Unidata Cloud Grants, Awards and Resources

Microsoft Azure Awards

Microsoft awarded two new \$20,000 "Azure for Research Grants" to Unidata. While these grants are time-limited, they provide us with invaluable resources to experiment with cloud computing environments. We are successfully running the THREDDS Data Server, RAMADDA, CloudIDV and EDEX servers in the Microsoft Azure Cloud.

Unidata has a new \$20,000 grant to install an EDEX server for community use, and to continue the testing and deployment of the CloudIDV. The EDEX server will become active in the coming month and will be announced to the community when it is available.

Progress has been made on the following:

- We have deployed numerous services and instances to the Azure Cloud, mirroring our experiments with the Amazon cloud infrastructure.
- We have been continuing to learn about Microsoft's Azure cloud infrastructure. In particular, Michael James and Julien Chastang have been experimenting with Virtual Networks (VNETs) for cost and time efficient data delivery via the LDM. Our initial foray into this area was somewhat disappointing as we concluded that Azure's VNET capability is not sufficiently mature. Specifically, our experiments for establishing VNETs across Azure subscriptions (so-called "peering") was not ready for the kind of LDM data transfers we had envisioned. This result means, at this time, it will not be

as easy to distribute data to our collaborators in the Azure cloud as we would have hoped for. We will check back in the coming months to see if Microsoft has made progress with peering.

XSEDE Jetstream Award

To further investigate how the Unidata community can benefit from Unidata technologies in the cloud, Unidata obtained an [XSEDE](#) equipment award on the [Jetstream cloud-computing platform](#). The Extreme Science and Engineering Discovery Environment (XSEDE) five-year, \$121-million award is a National Science Foundation supported project. We wish to continue our research of porting Unidata technology into a variety of cloud environments including non-commercial, research-oriented clouds such as Jetstream. Specifically, we would like to deploy a motherlode class machine on the Jetstream cloud with Docker technology in a manner similar to what we accomplished with our Azure resources. As Docker provides a common baseline for cloud computing, this experiment should proceed in a fairly smooth manner, but we will not know until we try. Jetstream became available in February of 2016. We are currently in the early stages of experimenting with Jetstream.

Progress has been made on the following:

We have done a small amount of testing on the XSEDE Jetstream Cloud. Our current allocation is too small to be useful. We will be requesting much more resources via an XSEDE Research Grant, shortly.

Dependencies, challenges, problems, and risks include:

Theoretically, the prototyping work that we have been doing on the Azure cloud should port smoothly to Jetstream because Docker technology promotes portability. However, we will not know until we try.

Amazon Awards

Progress has been made on the following:

- Learning about Amazon's cloud infrastructure
- Designing an initial architecture to support storing 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 to serve data stored on S3 in AWS.

Open Commons Consortium Award

The Open Science Data Cloud, a resource of the Open Commons Consortium (OCC), provides the scientific community with resources for storing, sharing, and analyzing terabyte and petabyte-scale scientific datasets. The OSDC is a data science ecosystem in which researchers can house and share their own scientific data, access complimentary

public datasets, build and share customized virtual machines with whatever tools necessary to analyze their data, and perform the analysis to answer their research questions. Unidata is a beta user of resources in the Open Science Data Cloud ecosystem and we have been provided cloud-computing resources on the Griffin cloud platform. Our allocations are renewed on a quarterly basis and Unidata is partnering with OCC on the NOAA Big Data Project. Given the limited staff resources and many ongoing cloud activities on AWS, Azure, and XSEDE environments, Unidata's activities on the OSDC have been in a temporary hiatus. We are hoping to ramp up our OSDC efforts in the upcoming months.

CloudIDV Application Streaming

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.

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, "CloudControl".
- We have released the new versions of the Dockerized IDV, CloudIDV.
- We have updated the generic application-streaming container for use by our community with their own legacy software, "CloudStream" to allow for encrypted-only and shared sessions.

AWIPS in the Cloud

The Azure for Research Grant for Unidata AWIPS has allowed for the edex-cloud open data server to live on, with an on-site EDEX server available as a replacement for those periods of time where a cloud-based server is not funded. This grant has enabled the development of a RedHat 7 supported EDEX and CAVE build, which can take advantage of the Azure file sharing architecture to create a distributed EDEX environment, scalable to data requirements.

A similar EDEX Data Server has been maintained by Unidata for Embry-Riddle Aeronautical University (ERAU) on an Amazon EC-2 instance, though access is restricted only to ERAU domains.

Both Azure and AWS have the capability to create machine images of these EDEX servers, further streamlining the process of installing a cloud-based EDEX server.

Progress has been made on the following:

- Staged development and testing EDEX data servers in the Azure cloud for both el6 and el7.
- Expanded the Azure EDEX data archive to retain 30 days of NEXRAD3 and FNEXRAD products.
- Submitted a talk to the 97th AMS Annual Meeting on the development and deployment of a distributed EDEX server in the Azure cloud.
- EDEX 16.2.2+ installs and leverages GEMPAK/nex2gini for real-time composite creation in the cloud (supplanting and expanding the FNEXRAD product set generated in AWS).
- Published [Jupyter notebooks](#) which provide access to real-time AWIPS data in the cloud, using the [python-awips](#) API.

Ongoing Activities

We plan to continue the following activities:

- Continue to create the LDM/IDD FNEXRAD (NEXRAD Level III national composite radar products) and UNIWISC (GOES-East/West image sectors) in a virtual machine hosted in AWS
- 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
- Maintain the TDS on AWS serving level II radar data
- Deploy Jupyter notebooks that provide access to NEXRAD Level II data stored in S3
- Deploy a distributed EDEX cluster in the cloud using a standalone database-class machine and Azure file share.
- Continue experimenting with our Azure resources for running Unidata technology in the cloud and staging motherlode class machine.

Big Data Project

- Continue to populate the NEXRAD level II archive with real time data.
- Continue to populate new GFS .25 degree output and NCEP HRRR output to an S3 bucket for access. We did not place a TDS on this collection as this output is available from our standard sources.
- TDS on AWS for level II NEXRAD: (For .edu access only)
<http://thredds-aws.unidata.ucar.edu/thredds/catalog.html>

- AWS Explorer: (Public access)
<https://s3.amazonaws.com/noaa-nexrad-level2/index.html>
- Public Bucket for level II NEXRAD: <https://noaa-nexrad-level2.s3.amazonaws.com>

New Activities

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

- Deploy new versions of CloudIDV and CloudStream
- Apply for an extension to our Microsoft Azure for Research award.
- Investigate the Jetstream cloud for running Unidata technology in the cloud and present the results at the 2017 AMS meeting in Seattle.
- Work with the Amazon Big Data Project team to bring GFS model output into the cloud in real time.
- Submit a new Microsoft Azure for Research proposal.
- Begin work on new Amazon Web Services grant to install and maintain an EDEX server in a cloud environment for the academic and research community to access
- Continue containerizing Unidata software and advocating for its use through presentations and blog entries.

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.
- Investigate cloud-based streaming services for CAVE deployment.

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

While Unidata is successfully moving its technology offerings to the cloud, we have not reinvented our technology to best take advantage of cloud computing. We hope to better research this area in the long-term.

Areas for Committee Feedback

- What clouds are our community using, either commercial or private?
- What new cloud technologies are our community using/investigating on their own initiative?
- Who would like to volunteer to beta test CloudIDV?

Relevant Metrics

Docker image downloads are available from [Unidata's Dockerhub repository](#).

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, EDEX, 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**
Transitioning Unidata technology to a cloud computing environment will increase data availability to new audiences thereby creating new and diverse geoscience communities.

Appendix

- [Cloud Computing Projects at Unidata Spreadsheet](#)
- [Cloud Computing Resources at Unidata Spreadsheet](#)

Status Report: Community Services

April 2016 - September 2016

Doug Dirks, Jeff Weber, Joshua Young, Larissa Gordon

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:

- [Python AWIPS Data Access Framework](#)
- [University of Wisconsin-Milwaukee Deploys AWIPS II](#)
- [2016 Community Equipment Awards](#)
- [Rosetta: A Data Transformation Tool for ASCII Files](#)
- [2016 Community Survey Results](#)
- [Western Kentucky University Students Dive Into AWIPS II](#)
- [Community Highlight: Millersville University and Their Involvement in the PECAN Field Study](#)
- [Successful Python Training Workshop in Madison, Wisconsin](#)
- [Unidata Program Center Welcomes Matt Perna](#)
- [Transitioning to the IDV and CAVE: Improving Classroom Technology for Meteorology at Plymouth State University](#)
- [Unidata Interns Wrap Up Summer Projects](#)
- 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:

- Assisting the Strategic Advisory Committee in their efforts to provide guidance for the next Unidata Strategic Plan.
- Continued work on the pilot project to assist community members with data management and document the resulting workflows for the DMRC was awarded

supplemental funding by NSF. Five academic volunteers are participating (Florida Institute of Technology, Lyndon State College, North Carolina State University, Millersville University, and the University of Wyoming). As a part of this project, community services intern Larissa Gordon joined the UPC through September 2016; she is finishing up her internship now.

- A Use Case and Scenario was created for Florida Institute of Technology documenting their data management strategy regarding a NOAA project they are a part of titled: An Ensemble-Based Approach to Forecasting Surf, Set-Up, and Surge in the Coastal Zone.
- A Use Case and Scenario was created for Lyndon State which documented their data management plan on a project titled Improving the Forecasting of Weather Driven Outages and Long Term Renewable Energy Production, which is seeking to improve predictions of climate change's effects on Vermont's renewable resources.
- A Use Case and Scenario was created for North Carolina State University documenting their data management plan for a project titled The Importance of Resolving Mesoscale Latent Heating in the North Atlantic Storm Track, which examines the effect of predicted increased latent heat from climate change on cyclogenesis.
- A Use Case and Scenario was written for Millersville University's data management strategy for their involvement in Plains Elevated Convection At Night (PECAN).
- A Use Case and Scenario was created for the data strategy of University of Wyoming's involvement in OWLeS (Ontario Winter Lake effect Systems) with their research aircraft King Air documenting atmospheric parameters surrounding lake effect systems.
- Present the DMRC at SciDataCon
- Work in providing a pilot online training (<http://unidata.github.io/online-python-training/>) focused on Python-content. This was the secondary project for community services intern Larissa Gordon
- Community engagement at recent professional society conferences (AMS Summer Meeting, Data Integrity Conference, Data Science as Narrative, International Coral Reef Symposium, International Data Week, Science of Team Science)
- Engagement with CUAHSI to support the NFIE and WRF-Hydro at the NWC
- Engagement with EarthCube as a Co-Chair of the Liaison Team
- Continue to serve on the CUAHSI HIS standing committee
- We continue to update Unidata's social media channels (Facebook, Twitter, Google+)
- We continue to publish short videos/screencasts on the [Unidata YouTube channel](#).
- We continue to actively support the NCAR/SOARS program
- Actively participate in Super Science Saturday
- Engage and support the Undergraduate Leadership Workshop (ULW) at UCAR
- Support the development and operation of the UCAR:NCAR Equity and Inclusion (UNEION) community of practice

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:

- Engagement with EarthCube, ESIP, and science or cyber communities at large
- 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
- Inclusion and equity
- 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
- Engage other UCAR/NCAR divisions regarding Unidata software use i.e. CESM/IDV

New Activities

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

- Expanded emphasis on cloud-related activities
- Seek additional resources to continue the online training resource focused on Python and Unidata services and tools
- Seek additional case studies for the DMRC with associated outreach through social media and the Unidata blog
- Support Committees work towards producing the next Strategic Plan

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

- Support Committees work towards producing the next Strategic Plan
- Continue to engage the hydrologic community regarding WRF-Hydro/IDV interactions and the National Water Center's efforts
- Seek additional opportunities to engage and listen to the community

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

- Provide additional data management and cloud-related training

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Please visit the [DMRC](#) and [OTP](#); these are both pilots but we would welcome your suggestions.

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:

- 52,386 unique pageviews (44,755 in previous period)
- 11.8% of total unique pageviews to site (7.9% in previous period)

Top community pages

1. All blog pages
39442 unique pageviews (33,825 in previous period)
75% of total community pageviews (76% in previous period)
2. www.unidata.ucar.edu/events/
6113 unique pageviews (3333 in previous period)
12% of total community pageviews (7% in previous period)
3. www.unidata.ucar.edu/about/
3316 unique pageviews (2800 in previous period)
6% of total community pageviews (6% in previous period)
4. www.unidata.ucar.edu/community/
2723 unique pageviews (4119 in previous period)
5% of total community pageviews (9% in previous period)

Social media statistics, March 31, 2016

1. # of Twitter followers: 552 (up from 503)
2. # of Facebook followers: 568 (up from 490)

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.

Status Report: Community Equipment Awards

Sponsored by the National Science Foundation

April 2016 - September 2016

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")
- Installation of a Jupyter notebook server in the cloud (e.g. JupyterHub or Wakari, but not limited to these) to experiment with the use of Python for cloud-based, data-proximate analysis and visualization

A Request for Proposals was sent out on December 2, 2015 with a March 25, 2016 submission deadline. The Review Panel met on April 12 at the Unidata Program Center and recommended that the following proposals be awarded:

- Oregon State University, Karen Shell, A Jupyter Server for the Oregon State University Climate Science Program, [Proposal](#)
- University of North Dakota, Gretchen Mullendore, UND Big Weather Web: Distributed Data Solutions, [Proposal](#)
- University of Louisiana at Monroe, Todd Murphy, AWIPS-II Infrastructure Upgrade at the University of Louisiana at Monroe, [Proposal](#)
- Pennsylvania State University, Chuck Pavloski, A Prototype Cloud-Based Visualization System for Unidata Applications, [Proposal](#)
- University of Iowa, Bong Chul Seo, A Pilot Project for Cloud-Based NEXRAD Data Processing, Analysis, and Visualization for Flood Forecasting and Water Resources

Management, [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 2017 Unidata Community Equipment Awards;**
- 2. Three volunteers from Users Comm to serve on the 2017 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 87 awards totaling close to \$1,100,000.

Prepared September 2016

Status Report: Internet Data Distribution

April 2016 - September 2016

Mike Schmidt, Jeff Weber, Steve Emmerson, Tom Yoksas

Activities Since the Last Status Report

Internet Data Distribution (IDD)

IDD data volumes continue to increase. The following output is from a Linux-based motherlode clone that the UPC operates on behalf of the community, lead.unidata.ucar.edu:

Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume 65031.001 M bytes/hour
Average hourly volume 38493.587 M bytes/hour

Average products per hour 407761 prods/hour

Feed	Average (M byte/hour)		Maximum (M byte/hour)	Products number/hour
FSL2	8715.140	[22.640%]	22132.535	14028.000
CONDUIT	8264.581	[21.470%]	23796.376	92980.682
NEXRAD2	7120.848	[18.499%]	9730.076	73132.364
NGRID	6038.497	[15.687%]	11252.577	39178.000
NOTHER	3527.528	[9.164%]	6238.838	10266.818
NEXRAD3	2331.997	[6.058%]	2973.654	100838.250
FNMOC	1284.306	[3.336%]	4039.717	3497.273
HDS	558.523	[1.451%]	917.453	27114.023
NIMAGE	224.196	[0.582%]	406.132	252.614
FNEXRAD	133.271	[0.346%]	166.668	106.682
GEM	129.827	[0.337%]	596.902	792.295
UNIWISC	72.877	[0.189%]	120.454	49.000
IDS DDPLUS	64.193	[0.167%]	72.514	44875.614
EXP	23.609	[0.061%]	46.660	266.614
LIGHTNING	4.066	[0.011%]	7.095	382.023
GPS	0.129	[0.000%]	1.165	1.068

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 (3) 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 receiving the NOAA/GSD data lose interest.

- The HRRR is being experimentally served at:
<http://thredds-jumbo.unidata.ucar.edu/thredds/modelsHrrr.html>
(.xml for machine access)
- Other data sets we continue to explore with NOAA/GSD/ESRL are:
 - [FIM](#)
 - [HIWPP](#)
- 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, 2015. The CMC is now feeding output of their new hi-resolution (15 km) GEM model to Unidata who, in turn, relays the data to IDD participants.

NOAAPort Data Ingest

- Ingest of the upgraded NOAAPort Satellite Broadcast Network (SBN) products and their relay to end-users via the IDD has been "operational" at the UPC since the August 2014.
- The NOAAPort-derived data streams (HDS, IDS|DDPLUS, NGRID, NIMAGE, NEXRAD3 and NOTHER) are being redundantly injected into the IDD at three geographically separate locations: Unidata, UW/SSEC, and LSU/SRCC. We are still seeking a fourth ingest site to increase robustness of the IDD distribution of NOAAPort derived data.
- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.13.4.
- Raytheon continues to submit modifications to the LDM for enhanced AWIPS functionality.

Relevant Metrics

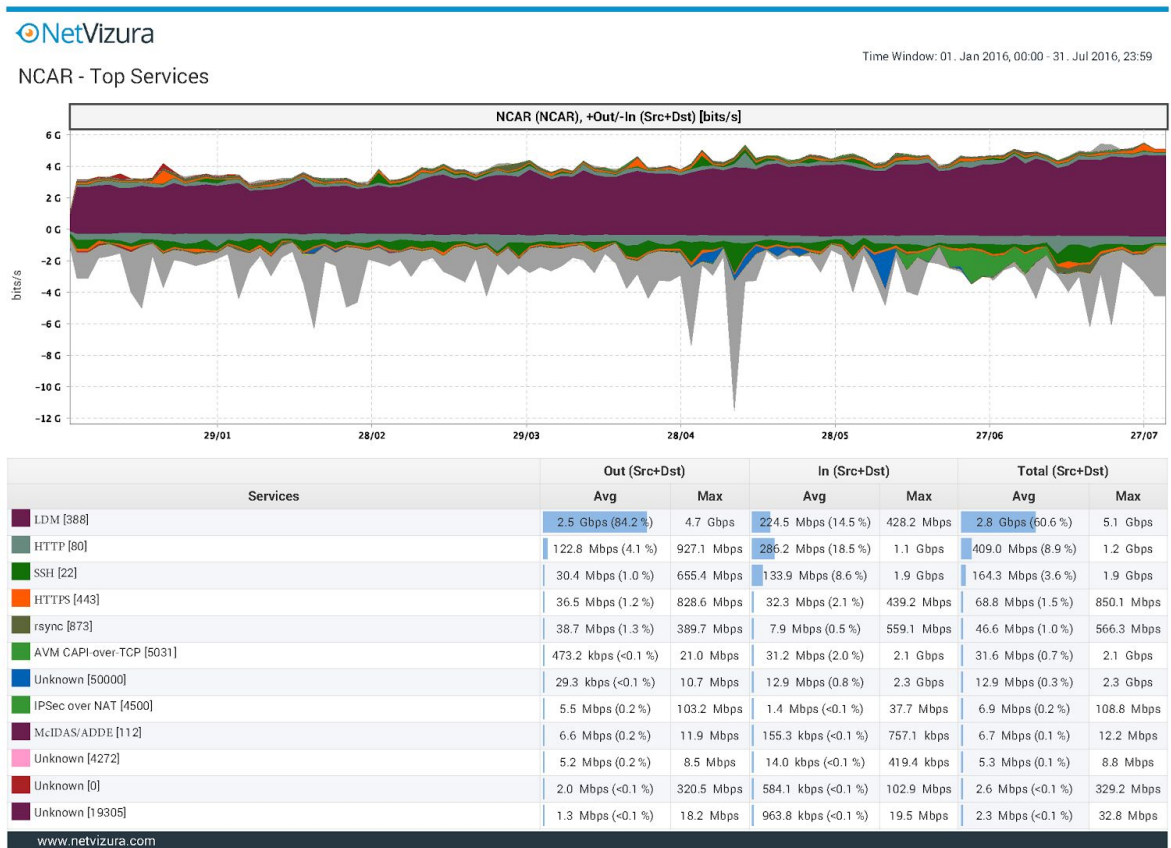
- Approximately **595** machines at **255** sites are running LDM-6 **and** reporting real time statistics to the UPC. Unidata staff routinely assist in the installation and tuning of

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 2150 downstream connections.

Data input to the cluster nodes averages around 29 GB/hr (~0.7 TB/day). Over the period from January 1 through July 31, the average volume of LDM/IDD data flowing in the UCAR/NCAR network exceeded 2.8 Gbps (~30 TB/day), and peak rates reached 5.1 Gbps (which would be ~55 TB/day if the rate was sustained).



Cluster real server backends and accumulator nodes routinely have instantaneous output volumes that exceed a Gbps. 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 IDD data volume over the past six months is largely attributable to the 0.25 degree GFS data in CONDUIT, the overall increase in the volume of data being transmitted in NOAAPort (which now routinely exceeds 12.5 GB/hr), and the increase in dual polarization NEXRAD data.

Questions for Committee Members

- Suggestions regarding content of datastreams like CONDUIT, FNEXRAD, UNIWISC and LIGHTNING? ... We (UPC, the Unidata community and UAlbany for the NLDN component of LIGHTNING) have control of the content of these datastreams, so they are open for suggestions.

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.

Status Report: IDV with RAMADDA

April-September 2016

Yuan Ho, Julien Chastang

Activities Since the Last Status Report

IDV System Changes

__Latest netCDF-Java Version__

The version of the netCDF-Java library currently distributed with the IDV is the 4.6.7 (April, 2016). [This was principally a bug fix release.](#)

__ISL Changes__

IDV Display Changes

__Latest Version of VisAD__

In the last six months, there have been several feature enhancements and bug fixes in VisAD:

- Volume vectors
- Volume trajectories
- Trajectories with different display types: line, ribbon, cylinder, and deform ribbon.
- Improved streamlines
- Introduced a workaround for Java 3D grey window/panel problem
- Added calibration to local McIDAS area files, currently only for the GVAR area file.

__Background Maps__

- The USGS shaded-relief background map is once again available.

__Other Changes__

- NetCDF point cloud display for WRF Hydro output is under development.

IDV WRF-Hydro Collaboration

The IDV team is collaborating with David Gochis from NCAR-RAL assisting him in ensuring WRF-Hydro data is CF compliant according Point Feature type. In addition, we are helping

David with visualization of this dataset the IDV.

IDV WRF-Tutorial Workshop

In the July WRF tutorial workshop, Yuan gave a 45 minutes presentation of the IDV main features and its applications in the WRF output datasets. We also help out the students in the classroom.

IDV Release

The last IDV release was [5.3](#) in May of 2016.

IDV EarthCube Proposal Awarded

In collaboration, with University of Miami Professor Brian Mapes, Unidata submitted an EarthCube proposal: "Drilling down from a statistics scatterplot to pre-populated case note books". This proposal was awarded \$1.3 million over three years.

IDV Publication Highlights

A [Google Scholar Search](#) reveals a number of publications that cite use of the IDV ([doi:10.5065/D6RN35XM](https://doi.org/10.5065/D6RN35XM)).

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 Docker container for easy distribution and installation of RAMADDA in a cloud environment. We coupled this effort with a Dockerized LDM with the goal of serving data in a cloud environment. We continue to work on this effort started last year and plan to present an update at the 2017 AMS annual meeting in Seattle Washington. In this example, [RAMADDA is serving IDD data supplied via the LDM](#).

IDV and RAMADDA Training, Conference Attendance and Presence

__2017 American Meteorological Conference (AMS) Annual Meeting__

- Planning IDV presence at the Unidata booth and student career fair

__2016 American Geophysical Union (AGU) Annual Meeting__

- Planning IDV presence at UCP booth

__2016 RMACC High Performance Computing Symposium

- [Cloud IDV demonstrated during Jetstream keynote](#)

__IDV Instructional Videos__

Posted a video on [volume trajectory display for Hurricane Sandy](#).

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.

__IDV Seam Issues__

We will continue to collaborate with the Unidata netCDF-Java team to devise a solution for long-standing longitude seam issues in the IDV.

New Activities

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

__IDV Training Workshop, October, 2016__

We will offer our annual training workshop in the Fall of 2016.

__Investigation of Java 3D Alternative__

Because of concerns about the long-term viability the open-source Java 3D project, the IDV team has begun discussions with our University of Wisconsin, SSEC collaborators to replace Java 3D with a more viable alternative within the VisAD API. We have started investigating whether the [Ardor 3D](#) can meet that objective. Looking into alternatives to Java 3D was also a goal described in the [Unidata 2018 Five-year plan](#).

Areas for Committee Feedback

We have no questions at this time.

Relevant Metrics

__E-Support__

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 ~70 e-support tickets. Each individual ticket may and often does involve many back-and-forth messages. There is an especially large number of support requests coming from international users.

__Usage Metrics__

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

Top ten universities running IDV are: Millersville, Oklahoma, University of Utah, St Cloud state, Plymouth, NC State, West Kentucky, Lyndon State, University of Illinois, and San Francisco State.

__Github Pull Requests__

In the area of greater collaborative development, since the migration of the IDV project to github, we have closed a total of 100 “pull requests” or code contributions from internal and external collaborators.

__Youtube IDV Instructional Videos__

In the area of online IDV training, the Youtube IDV instructional videos have been viewed 10,000 times compared with 8,000 from six months ago.

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. One of RAMADDA's best features is the ability to upload a CDM file and obtain the OpenDAP link from the new entry. The DAP link can be shared and opened in the IDV. RAMADDA also has "server-side view" capability where a specific part of the file system can be made available through the RAMADDA interface. This feature can be helpful to view LDM data feeds, for example. 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.

Status Report: International Activities

April 2016 - September 2016

Tom Yoksas, Yuan Ho

Activities Since the Last Status Report

Presentations at Guangdong Ocean University and Beijing China FAA

Yuan gave impromptu talks at Guangdong Ocean University and Beijing China FAA meteorological office in June. The presentations focused on the IDV and the Common Data Model as it relates to Chinese level2 radar data.

Presentations, Tutorials and meetings at the Chinese Meteorological Administration main office in Beijing and provincial office in Xining

Unidata conducted a multi-day training event at the Chinese Meteorological Agency offices in Beijing, China on August 13-26.

The tutorials focused on the use of the LDM to distribute meteorological data, the IDV to visualize data, and future directions in satellite imagery.

New/Planned Activities

No new activities are currently planned

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

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.

Prepared *September, 2016*

Status Report: LDM

April 2016 - September 2016

Steve Emmerson, Tom Yoksas, Mike Schmidt, Julien Chastang

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:

- Made logging more robust and improved its performance
- Made the handling of the standard output and standard error streams more robust in relationship to child processes
- Merged and (corrected) Raytheon's use of zlib(3)
- Improved the performance of pqact(1)
- Made the product-queue more robust
- Corrected the handling of blank but non-empty lines in the LDM configuration-file
- Ported improvements to OS X and Solaris
- Improved the documentation
- Released versions 6.13.1, 6.13.2, 6.13.3, and 6.13.4
- The LDM package was successfully containerized using Docker technology. See the [Cloud Status Report](#) for more information.

Dependencies, challenges, problems, and risks include:

- Dealing with missing GEMPAK table entries represents a problem area
- The LDM is sometimes held responsible for decisions made by the NWS on how to categorize data products.

Multicast LDM (aka 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:

- Testing of LDM-7 over static VPN circuits on Internet2 was successful. Participants included the University of Virginia, Rutgers University, Indiana University, and the University of Wisconsin.

Dependencies, challenges, problems, and risks include:

- NSF funding the new deployment proposal
- Time

Ongoing Activities

We plan to continue the following activities:

- Support and maintenance of the LDM
- Continue adapting the LDM to the Docker container technology to support cloud activities and (perhaps) make life easier for LDM users

New Activities

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

- Development and deployment of LDM-7 (assuming the proposal is funded)
- Conduct an LDM training workshop for NOAA/NWS at College Park

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

- Development and deployment of LDM-7 (assuming the new proposal is funded)

Relevant Metrics

- Data on the LDM package can be found [here](#)
- The LDM system at the Unidata Program Center powers the IDD. Metrics on that program can be found in the IDD status report.

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 and developing multicast technologies.
-

Prepared *September, 2016*

Status Report: McIDAS

April 2016 - September 2016

Tom Yoksas

Activities Since the Last Status Report

Aside from routine updates/bugfixes to existing code and tables, the main area of investigation has to been to add indexing to ADDE datasets to speed access into large datasets.

Prior Activities

- Unidata McIDAS version 2016 was made available in March. v2016 includes all SSEC versions up to and including the current release, v2016.1 and Unidata updates and bugfixes.

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

The latest releases features the following:

- Introduction of and an update to an ADDE server for Himawari imagery (this is a precursor for a GOES-R ADDE server)
- Initial release of an ADDE server for GOES-R imagery. It is believed that this server will not work with the GOES-R image sectors that will be included in the NOAAPort SBN. If this is, in fact, the case, new development will be needed to support those images.

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-X and help using McIDAS-X 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

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-X/-XCD 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 routinely exceeds 2 TB/month.
- Direct, programmatic access to real-time GOES-West (GOES-15) data via McIDAS ADDE routinely exceeds 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. Launch of GOES-R is currently planned for November 4, 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 way as Landsat imagery.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

- 1. Enable widespread, efficient access to geoscience data**
*MclIDAS remains **the** application of choice for the satellite meteorology community. The Advanced Data Distribution Environment (ADDE) component of MclIDAS 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 MclIDAS, MclIDAS-V, unlike its predecessors, is a fully open source application that is in wide scale and growing use in the satellite meteorology community. MclIDAS 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**
MclIDAS is sought for use by those interested in satellite meteorology worldwide.
-

Prepared *September, 2016*

Status Report: netCDF

April 2016 - September 2016

Ward Fisher, Dennis Heimbigner

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 29 open issues for netCDF-C, 13 open issues for netCDF-Fortran, and 7 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:

- Further extension of the netCDF build-and-test platforms using Docker technology.
- Further enhancements to the netCDF documentation.
- Extended continuous integration platforms.
- Migrating away from the EOL JIRA platform to the GitHub issue tracking dashboard for issue and feature-request tracking.

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.
- The hdf5 1.10 version generated, by default, backwards-incompatible binary netCDF4 files. This was addressable but was a short-notice high-priority issue which required immediate attention.

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:

- Seek out, and prepare material for upcoming, conferences and other outreach opportunities.
- Work on reducing the defects reported by static analysis.
- 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:

- Continue 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. **Should netCDF be ported to and/or maintained for any other programming computing/development environments?**
2. **Are there any emergent avenues (stack overflow, etc) for user support which the netCDF team should investigate?**
3. **How can we encourage more user testing of the release candidates we provide?**

Relevant Metrics

There are currently about 142,810 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 increased slightly from **0.32** six months ago to **0.36** 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:

- **652,000** for netCDF-3
- **484,000** for netCDF-4
- **373,000** for HDF5
- **61,800** for GRIB2

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

- **246** for netCDF-3
- **435** for netCDF-4
- **7,230** for HDF5
- **617** for GRIB2

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.

Prepared *September 2015*

Status Report: Outreach Activities

April 2016 - September 2016

Ben Domenico

Activities Since the Last Status Report

[Open Geospatial Consortium \(OGC\) and Ocean Data Interoperability Platform \(ODIP\) activities](#)

Continue to work with OGC to augment international CF-netCDF standards that have been established over the last several years and on phase 2 of ODIP.

Progress has been made on the following:

- Represented Unidata at OGC Technical Committee meetings
- Chaired [OGC NetCDF Standards](#) Working Group
- Served as UCAR business and technical representative to the OGC.
- Represented Unidata at ODIP Steering Committee telecons
- Acted as Unidata host for [ODIP Workshop in May 2016](#)
- Started transfer of responsibility to other staff for ongoing projects.

Dependencies, challenges, problems, and risks include:

- There is a question of what resources should be committed to these projects beyond the beginning of the 2017 calendar year when Ben retires from Unidata.

[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:

- Maintained periodic contact with Liping Di, the GMU PI on the project.
- Did mid-course correction to objectives based on one year no cost extension to the project.

Dependencies, challenges, problems, and risks include:

- The one year extension enables us to use the remaining resources available in the project to accomplish tasks that are in line with the strategic goals for the new version of the TDS and to ensure that our data are available via standard interfaces such as those provided by the Data Discovery and Access Broker of the Italian CNR group Python Workshop Notebooks on alternative platforms.

Due to reduced FTE commitment, no effort is being expended on cloud software development

Ongoing Activities

- Coordination and collaboration with [NCAR GIS](#)
- Represent Unidata and UCAR in OGC
- Participate in ODIP 2 as resources allow after hosting workshop in May
- Continue as co-PI on [Earthcube Cyberconnector project](#).

Note that the responsibility for these ongoing activities is being transferred to other staff.

New Activities

No new activities are planned. For the remainder of this calendar year, the objective is to keep current activities going as well as possible with resources available and make arrangements to complete those that can be completed. Note that, in response to last year's budget crunch, Ben reduced his Unidata FTE commitment.

Areas for Committee Feedback

Suggestions for the transfer of responsibilities to other staff members are welcome.

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.

Status Report: Python

April 2016 - September 2016

Ryan May, Sean Arms, Julien Chastang, Ward Fisher

Activities Since the Last Status Report

Python Training Efforts

- We continue to improve and expand the Unidata Python Workshop, ensuring it stays current with the latest development in the scientific Python area. Specifically, we enhance and add Jupyter Notebooks in the geoscientific domain in addition to making sure we stay on top of infrastructure changes with the conda package manager, and the Jupyter environment.
- Unidata obtained supplemental funds from NSF to start an [online Python training](#) effort specifically focused on serving the geoscience community. This was a solid pilot effort towards creating online Python training materials, but much remains to be done before it is a complete resource.
- We have started fostering a GitHub repository for collecting useful example notebooks for using Python in the geosciences. Kristen Pozsonyi, one of Unidata's summer interns, helped spearhead an effort to turn these notebooks into a much better [online gallery](#). We welcome community contributions to this repository.
- Sean Arms and Ryan May travelled to Madison, WI to present a pair of two-day workshops on Python for CIMSS/SSEC and the University of Wisconsin Atmospheric and Oceanic Sciences Department. The workshop had over 50 attendees, and the feedback for the workshop was overwhelmingly positive. Lessons learned from these workshops is being used to improve the annual Python training materials.
- Thanks to the tutorial sessions at SciPy 2016, Ryan May is now a certified instructor for Software Carpentry (SWC). This led to his travelling to Texas Tech University to help give a SWC workshop sponsored in part by the TTU Atmospheric Science group. This was an excellent opportunity to teach git and introductory Python to a diverse group, which again has provided lessons on how to improve our annual Python workshop. This was also an opportunity to introduce TTU Atmospheric Science graduate students to MetPy, etc., as well as solicit contributions.

Progress has been made on the following:

- Creating introductory online Python training materials
- Continued improvement and refinement of the annual Python workshop materials--to the extent that there is less individual preparation for each workshop.
- Expansion of available training materials and in-person training offerings
- Creating a repository for gathering Jupyter notebooks created both internally and by the community as a learning resource

Dependencies, challenges, problems, and risks include:

- The amount of training has contributed to a slower pace of development on MetPy/Siphon--though this is by no means necessarily a bad thing.

MetPy

The MetPy community continues to grow slowly. There have been several externally driven Pull Requests, both for bug fixes and new features; the MetPy [twitter account](#) has also reached 105 followers. Both of Unidata's summer interns, Alex Haberlie and Kristen Pozsonyi, spent the majority of their time working on additions to MetPy. Kristen's time was especially well-spent; as a beginner to Python, the challenges she experienced have helped inform some areas where MetPy's user experience needs improvement.

We also continue to improve MetPy's open development model. Ideas for further development are captured in the GitHub issue tracker; current plans are to focus on enhancing the abilities to read point data (both BUFR and raw METAR) to facilitate generating NetCDF data for hosting on THREDDS, as well as simplifications to MetPy for working with units. The MetPy project is investigating some development changes, such as greater use of milestones for planning, a time-based release schedule, and active solicitation of feedback on these milestones in consideration of additional staff hired to work on Python, and to facilitate more community feedback in priorities.

Progress has been made on the following:

- Continued investment in MetPy's automated build, testing, documentation, and release infrastructure
- Kristen Pozsonyi added more calculations to MetPy
- MetPy now has capabilities (still considered beta-level in terms of API) for interpolating point data to a grid using a variety of methods, such as Barnes, Cressman, and Natural Neighbor (Thanks, Alex Haberlie!)
- The MetPy examples now show up as a gallery of images thanks to Kristen Pozsonyi's hard work
- A calculation for sounding equilibrium layer (and curve intersection algorithm) was contributed by the community
- Community awareness and involvement progressing well one year into the project

Siphon

[Siphon](#) represents our official Python support for TDS. While development has been slow of late, this is largely because its current capabilities meet current needs (versus some needs in MetPy). We anticipate developing Siphon to ensure that it is as easy as possible to download data from a TDS in Python (such as crawling a server looking for data), keeping pace with new features added on the Java side.

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 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. Ryan May is also now a member of the planning committee for the Python Symposium at the AMS Annual Meeting; he will also be presenting the Core Science Keynote for Python at the annual meeting, as well as helping bring the future generation of students into fold by presenting on Python at the AMS Student Conference.

Progress has been made on the following:

- Have continued to evaluate xarray (formerly xray, created by Stephen Hoyer) as a way to get CDM-like functionality in Python. It's current abilities provide a nice coordinate-aware data-object, as well as a way to attach attributes to arrays. This project has become a general tool for scientific Python, elevated to a top-level project within [PyData](#).
- Participated in the [conda-forge](#) project on GitHub; this is a community project to produce automated builds of conda packages using open recipes and infrastructure. We have contributed (and maintain) recipes for MetPy and Siphon, as well as their dependencies. These packages are available from the conda-forge channel on [anaconda.org](#)

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
- Growing and developing MetPy as a community resource for Python in meteorology
- Continued participation in the scientific Python community
- Relevant matplotlib support and fixes
- Working with JupyterHub as a way to facilitate data-proximate analysis
- Continue regular series of notebook-based blog posts on the Unidata Developer's blog to demonstrate the use of Python for various meteorological tasks
- As resources and time permits, continue making progress on the Online Python Training project by writing Jupyter notebooks specifically targeted towards teaching the geoscience community programming concepts. We have submitted an abstract to present this project at the AMS 2017 Annual Meeting in Seattle, WA.

New Activities

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

- Unidata's Annual Training workshop on "Using Python with Unidata Technologies"
- Help hold AMS short course on accessing the NEXRAD archive in AWS at the 2017 AMS Annual Meeting

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

- Restructure our annual Python workshop to be a full week with introduction to Python/git, intermediate with MetPy/Siphon/etc., and developer hack-day
- Attend SciPy 2017
- Python related presentations by Unidata staff at the 2017 AMS Annual Meeting in Seattle, WA.

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
- Offer a version of our Python workshop as an AMS short course at the Annual Meeting in 2018

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. In light of the success of the Madison Python workshop, should we be offering more regional Python workshops? Is anyone willing to help sponsor?
2. Does offering our training at AMS (or other conferences) seem like a worthwhile effort in order to take more advantage of the opportunity presented by a gathering of our community? Are there other conferences in addition to AMS you would suggest?
3. Are there any additions you'd like to make to MetPy's or Siphon's development roadmap?
4. What are the biggest obstacles that you see to the use of Python with other Unidata technologies, or for use in meteorology in general?
5. We continue to maintain the Unidata Python Workshop with fresh, relevant, and up-to-date content. However, we would welcome feedback from our committees on topics we may not be covering in the workshop.

Relevant Metrics

Siphon

- 95% test coverage
- 791 downloads/month from the Python package index
- Watchers: 10
- Since 1 April 2016:
 - Active Issues: 14 (5 created, 3 closed)
 - Active PRs: 8 (7 created, 8 closed)
 - External Issue Activity: 0 opened, 0 comments
 - External PR Activity: 1 opened, 1 comments
 - Unique external contributors: 1

- Stars: 5 (26 total)
- Commits: 31
- Since 1 October 2015
 - Active Issues: 37 (24 created, 15 closed)
 - Active PRs: 30 (28 created, 30 closed)
 - External Issue Activity: 1 opened, 7 comments
 - External PR Activity: 1 opened, 5 comments
 - Unique external contributors: 6
 - Stars: 12 (26 total)
 - Commits: 169

MetPy

- 94% test coverage
- 1670 downloads/month from the Python package index
- Watchers: 22
- Since 1 April 2016
 - Active Issues: 63 (50 created, 22 closed)
 - Active PRs: 43 (41 created, 36 closed)
 - External Issue Activity: 24 opened, 34 comments
 - External PR Activity: 20 opened, 20 comments
 - Unique external contributors: 19
 - Stars: 22 (81 total)
 - Commits: 135
- Since 1 October 2015
 - Active Issues: 103 (70 created, 41 closed)
 - Active PRs: 70 (69 created, 63 closed)
 - External Issue Activity: 27 opened, 45 comments
 - External PR Activity: 23 opened, 21 comments
 - Unique external contributors: 24
 - Stars: 38 (81 total)
 - Commits: 288

Unidata Python Workshop

- Watchers: 24
- Since 1 April 2016
 - Active Issues: 16 (14 created, 4 closed)
 - Active PRs: 17 (17 created, 17 closed)
 - External Issue Activity: 1 opened, 3 comments
 - External PR Activity: 0 opened, 0 comments
 - Unique external contributors: 1
 - Stars: 15 (39 total)
 - Commits: 64
- Since 1 October 2015
 - Active Issues: 23 (21 created, 7 closed)
 - Active PRs: 23 (23 created, 23 closed)
 - External Issue Activity: 1 opened, 3 comments

- External PR Activity: 0 opened, 0 comments
- Unique external contributors: 1
- Stars: 19 (39 total)
- Commits: 98

Unidata Online Python Training

- Watchers: 4
- Since 1 April 2016
 - Active Issues: 31 (21 created, 26 closed)
 - Active PRs: 44 (44 created, 44 closed)
 - External Issue Activity: 0 opened, 2 comments
 - External PR Activity: 6 opened, 1 comments
 - Unique external contributors: 2
 - Stars: 0 (0 total)
 - Commits: 76
- Since 1 October 2015
 - Active Issues: 40 (40 created, 26 closed)
 - Active PRs: 49 (49 created, 49 closed)
 - External Issue Activity: 0 opened, 2 comments
 - External PR Activity: 6 opened, 1 comments
 - Unique external contributors: 2
 - Stars: 0 (0 total)
 - Commits: 89

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. 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 the HTML5 Jupyter Notebook technology has the potential to address "very large datasets" problems. Jupyter Notebooks can be 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 and the goals outlined [Unidata 2018 Five-year plan](#). 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. The additional development of [MetPy](#) fills the need for domain-specific analysis and visualization tools in Python.

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.

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, now hosts a three day session to explore [Python with Unidata technology](#). Also, to advance the use of NetCDF in Python, Unidata has promoted Jeff Whitaker's [NetCDF4-python project](#), including hosting its repository under Unidata's GitHub account. Unidata is initiating a project to provide [online Python training](#) specifically targeting geoscience students. Unidata is also fostering some community development of meteorology-specific tools under the MetPy project.

Status Report: Support

April 2016 - September 2016

Tom Yoksas, Jennifer Oxelson, UPC Staff

Activities Since the Last Status Report

Training

- The UPC will host its annual training workshop series from October 17 - November 4, 2016 in its Foothills Lab offices in Boulder, CO.

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 50500 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., github, Stack Overflow, mailing list replies, etc.) add substantially to the support load.

Additional metrics may be found [here](#).

Support by Category

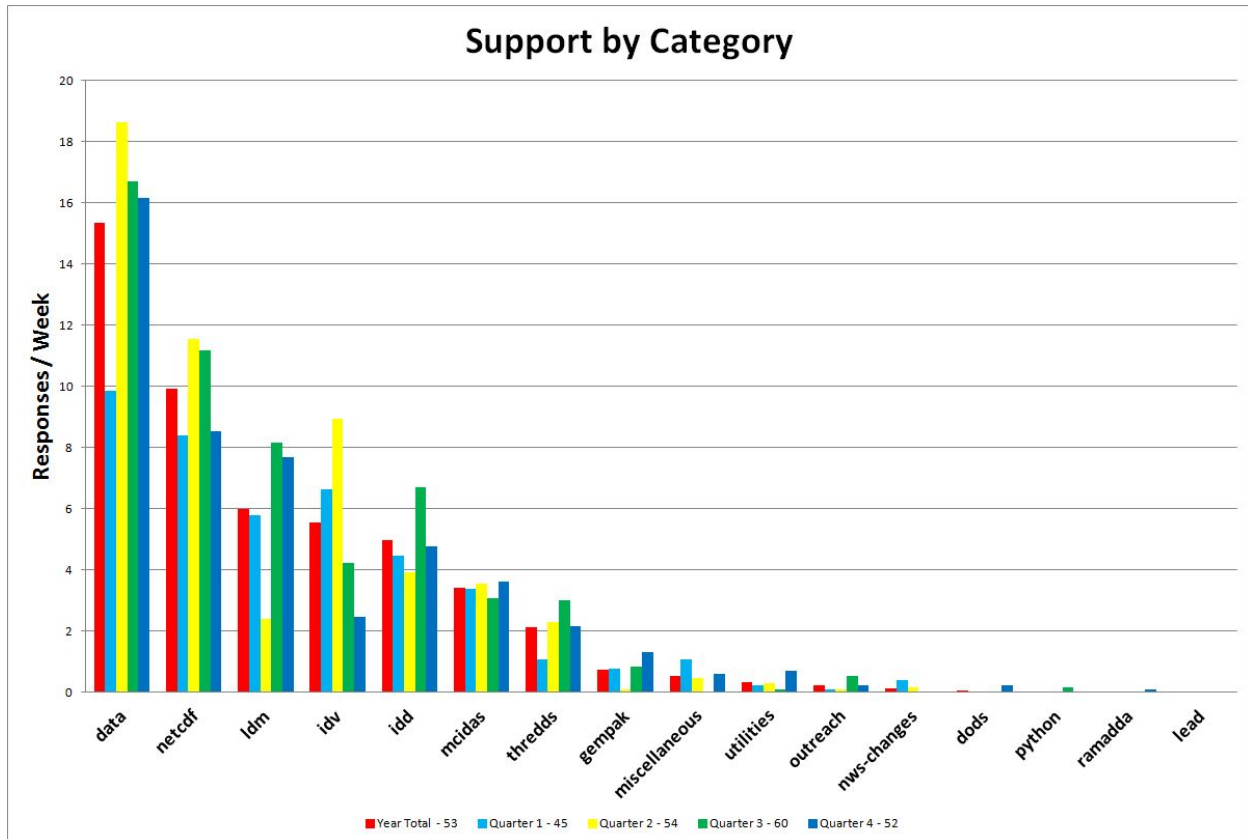


Fig. 1: Above are histograms that portray the number of Unidata email responses for categories of support for a one year period ending September, 16, 2016. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period 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, 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

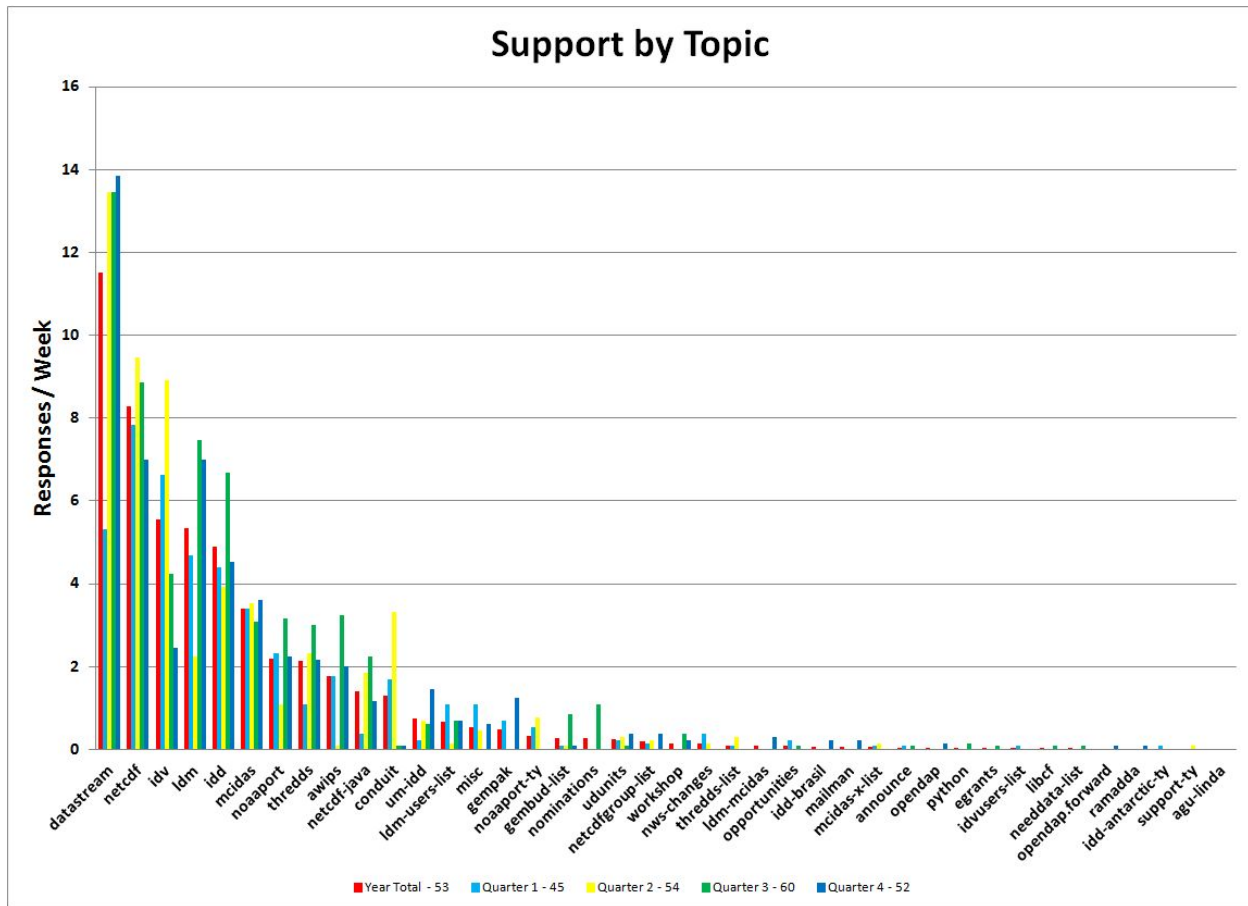


Fig. 2: Above are histograms that portray the number of Unidata email responses for individual topics of support for a one year period ending September 16, 2016. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the 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.
- 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 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.
 - 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](#)

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.

Prepared *September, 2016*

Status Report: THREDDS

April 2016 - September 2016

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

Activities Since the Last Status Report

The THREDDS Project

The THREDDS Project encompasses four projects: netCDF-Java, the THREDDS Data Server (TDS), Rosetta, and Siphon (the Unidata Python client to interact with a TDS). For specific information on Siphon, please see the Python Status Report. An update regarding cloud efforts related to the TDS can be found in the Cloud Computing Activities Status Report.

Action Item Follow-up from the Spring 2016 Meeting

1. *ACTION 10: Distribute THREDDS pqacts*

The pqact files containing all of pattern actions used to drive the Unidata TDS (<http://thredds.ucar.edu/thredds/catalog.html>) can be found on github (<https://github.com/Unidata/TdsConfig/tree/master/idd/pqacts>)

2. *ACTION 2: 84 servers running TDS, UPC will put together a list for future distribution to the committees*

The list of reachable TDS is now up to 172. Of the 172 reachable ip addresses, 133 resolved to an active url. See the list of active urls [here](#)

Released netCDF-Java / TDS version 4.6.6 (Stable)

Progress has been made on the following:

- The 4.6.x line of development is now in maintenance mode so that the team can focus on v5.0

Focus netCDF-Java / TDS (Soon-to-be Beta) v5

The THREDDS team is preparing to release a beta version of the THREDDS Data Server version 5.0, set for the beginning of November

Progress has been made on the following:

- New Coverage data type allows for subsetting across array boundaries (often called the "seam" problem)
- Uses the new edal-java based ncWMS 2.0 server, as well as javascript client Godiva3

- CatalogScan feature allows for incremental updating of TDS catalogs without the need to restart Tomcat
- Upload/Download support has been added to TDS.
- Unit and Integration tests are passing in 5.0. This is a big step towards releasing a beta.

Dependencies, challenges, problems, and risks include:

- The longer the 4.6.x line of development is maintained, the longer it will take to move forward with the 5.x line of development
- John Caron is now employed by Google, and as such will have minimal to no involvement with future TDS development. Prior to his employment with Google, John was going to continue to extend TDS on a contract basis.

Rosetta

The THREDDS team has extended the transformation capabilities of Rosetta.

Progress has been made on the following:

- Rosetta now supports the Trajectory DSG, which means CSV datasets from aircraft, drifting buoys, radiosondes, etc. can now be transformed into netCDF CF-1.6 compliant files.
- Rosetta now support automatic translation of the "EOL Sounding Composite" ASCII-based format into netCDF CF-1.6 compliant files. This capability was added to support the Data Management Resource Center's work with Millersville University, in which all radiosonde datasets from the PECAN project are being transformed into CF compliant netCDF files.

Dependencies, challenges, problems, and risks include:

- Two of the core javascript libraries used by Rosetta have been abandoned by their original creators. One has been picked up by the community (SlickGrid), while the other is in limbo (jWizzard). Unidata will likely need to pick up jWizzard and maintain it for use within Rosetta, at least internally. However, it would be a good community service to open this up to a wider audience, but resources would be required to do so.

Ongoing Activities

We plan to continue the following activities:

- Documentation updates - We are reworking the tutorial material for the TDS v5.0 with the goal of enabling asynchronous training. The material will undergo a major overhaul to include the use Docker containers, video snippets, and other new forms of training tools.

- 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>) Only the profile DSG is left to implement.

The following active proposals directly involve THREDDS work:

- 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.
- [NEW] A NASA ROSES ACCESS proposal has been funded and started at the beginning of September. The award is titled: "Leveraging available Technologies for Improved interoperability and visualization of Remote Sensing and in-situ Oceanographic Data at the PO.DAAC" and was submitted with JPL/PO.DAAC. [Rosetta]

New Activities

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

- Releasing a beta of TDS 5.0
- Officially advertizing a public TDS 5.0 Test Server [currently found at <http://thredds-test.unidata.ucar.edu/thredds/catalog.html>]
- Finalize visualization preview of converted data in Rosetta

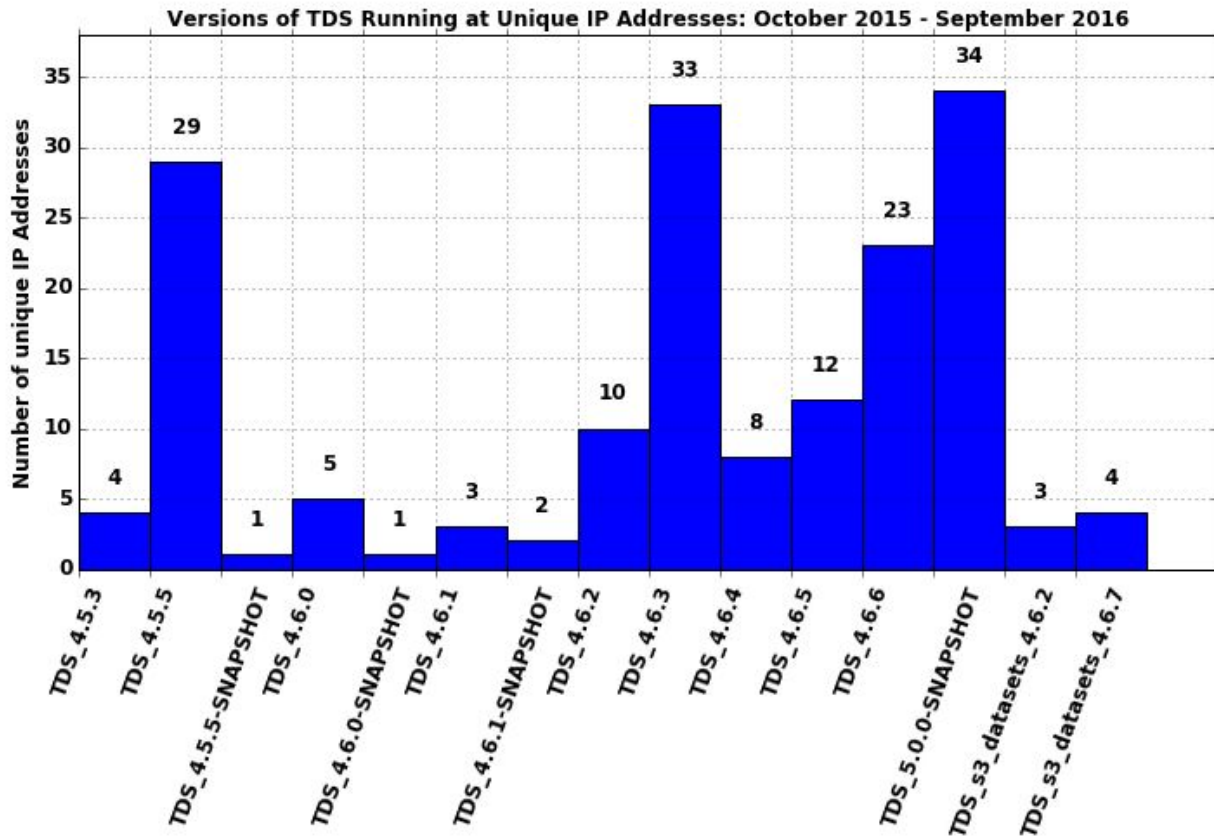
Over the next twelve months, we plan to organize or take part in the following:

- Create a TDS plugin layer for external services
 - Finalize the TDS plugin layer.
 - Upgrade the ncWMS, ncISO, and other plugin services to use the new TDS 5.x plugin layer
 - Incorporate ncSoS into TDS
- Transitioning thredds.ucar.edu to TDS 5.x
- Getting TDS v5.0 to a stable release
- Getting netCDF-Java v5.x to a stable release

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

- Enable Rosetta to publish to a TDS

Relevant Metrics



2,410 unique IPs started up threads between October 2015 and September 2016, 172 of which are publicly accessible servers. This information is only known for servers running v4.5.3 and above. The differences in these numbers could be due to:

- Reporting TDS running behind a firewall that does not allow incoming traffic on 80 or 8080 (the ports tested)
- It might be possible that a TDS running through a proxy server may not be “seen” in this analysis as publicly reachable
- People testing the TDS on their local machine, but not actually running a server

Note that the vast majority of the publicly accessible servers are running v4.6.3 or above (v4.6.6 was the most current release during this period, and was released on 16 June 2016). This indicates that users and organizations running the TDS tend to follow along closely with the current releases of the TDS.

This analysis also indicates a number of sites are running TDS v5.0, even though it is pre-beta. This underscores the desire for the new features in 5.0, and highlights one reason why we feel the need focus most of our efforts on and to move all new development to the

v5 codebase.

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.

Prepared *September 2016*