Agenda

Monday, 11 April 2016

8:30-9:00 Continental breakfast

9:00 - 9:15 Welcome and Administrative Items - Russ Schumacher

- Review of Action Items
- Date for Fall meeting

9:15 - 10:00 Directors Report - Mohan Ramamurthy

10:00 - 10:15 Break

10:15-10:45 SAC Report

10:45 - 11:45 CONDUIT/AWIPS II Becky Cosgrove

11:45 - 1:00 Lunch

1:00 - 1:30 Python Ryan May

1:30 - 2:45 Community Survey Results/Strategic Planning Process Session 1

Overview Presentation (JY)

2:45-3:00 Break

3:00-4:45 Community Survey Results/Strategic Planning Process Session 2

4:45 - 5:00 All Other Business

Time TBD Group Dinner

Tuesday, 12 April 2016

8:30-9:00 Continental breakfast

9:00 - 9:40 Around the Table Reports/Blue Skies

9:40-10:00 IDV Turbulence Demonstration (Yuan Ho)

10:00-10:15 Break

10:15-10:35 DeSouza Nominees Discussion

10:35-11:15 Status Report Discussion

11:15–11:35 IDD/LDM Updates (Steve Emmerson & Tom Yoksas)

11:35- 12:00 All Other Business

12:00 Adjourn

Status Report: ACADIS

September 2015 - April 2016

Mohan Ramamurthy, Sean Arms, Jeff Weber

Activities Since the Last Status Report

The Rosetta project

After eight successful years of the ACADIS project at UCAR and NSIDC, the NSF Arctic Data Service award was made to University of California, Santa Barbara. As a result, the ACADIS team is currently is working with the UC Sanata Barbara team in transitioning the data holdings to their system.

Unidata's main contribution to the ACADIS project was the creation of Rosetta, a web-based data translator. The Rosetta project is now being maintained and developed under the THREDDS project umbrella within Unidata.

While the ACADIS workflow for Rosetta is no longer in service, access to Rosetta is still provided at http://rosetta.unidata.ucar.edu.

Since the ACADIS project has ended, this will be the final status report on ACADIS.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

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

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

---

Status Report: AWIPS II and GEMPAK

September 2015 - April 2016

Michael James

AWIPS II Activities Since the Last Status Report

Python Data Access Framework (python-awips)

In March 2016 I published a standalone Python Data Access Framework (DAF) for retrieving data from an AWIPS II 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.

- The DAF was released as its own Python package and is available on PyPI:
  - **pip install python-awips**
- Bundled in the AWIPS II Python Stack but runs on its own as well:
  - **yum install awips2-python-awips**
- Documentation hosted at http://python-awips.readthedocs.org/en/latest/
  - Integrated publishing triggered by push to github repo
  - **Gridded Data** plotted with **Basemap**
  - **Gridded Data** plotted with **Cartopy**
  - **NEXRAD Level 3** plotted with **Matplotlib**
  - **Upper Air** sounding with **MetPy**
  - **Surface Obs** plot with **MetPy**
  - **Synop**, **Marine**, **Profiler**, **ACARS**, **AIREP/PIREP** data
  - Developer's guide for a detailed overview of the technology.
- Published a set of Jupyter/IPython notebooks with data plotting examples.
- This library is also used by GEMPAK to retrieve data from a remote EDEX Data Server, with the AWIPS model name defined in **datatype.tbl**

AWIPS II User Guide on Github Pages

This new AWIPS II User Guide on Github Pages expands on the original Unidata AWIPS II workshop documents and provides data visualization examples for a growing set of CAVE plugins.

For the AWIPS II User Guide, I published a new framework for building documentation on Github Pages. Polymer Web Components, Jekyll, Vulcanizer, and Grunt are used to publish rich and dynamic Markdown-based documents.

AWIPS II Release Notes

15.1.1, Jan 21, 2016

- First OS X CAVE client
- Moved from Java 1.6 to 1.7
- OS X client required openGL bindings be refactored from JOGL 1 to JogAmp 2
- Added install scripts to manage EDEX and CAVE packages (config, install, update)
- Various RPM install/uninstall bugs
- Automatically reset purgejobs on startup (which was preventing EDEX from scouring data and thus filled up disks)
- MRMS grids support both from NOAAport and via direct feed from NSSL
- Various GFE client fixes (muted vcmodule to allow the GFE client to launch)
- Disabled all **nwsauth**, **useradmin**, and **archive** plugins
- **Qpid** updated to 0.32 to solve restart problems
- New colormaps for GOES sounder McIDAS product CAVE displays
- Muted all **LDAD** components
- Tested and confirmed **archive** plugins, muted for build (to prevent disks from filling)
- Image export now includes timestamp in the filename by default (to avoid repeated overwriting prompt for default filename)
- Incorporated a postgres radar table update into the execution of an RPM update to avoid users having to uninstall/reinstall their database.

15.1.2, released Feb 23, 2016

- OS X client now runs off its own bundled JRE
- Built libgempak and libnsharp for OS X distribution
- Single file sourced from /etc/profile.d/ rather than multiple files.
- ECMWF global grid compositing (replacing the previous N Hemisphere-only compositing)
- CAVE menubar buttons to open NCP/GEMPAK-like data displays from within D2D
- New satellite imagery in menus from UNIWISC feed
- New grid bundles for default models, ocean models, and mesoscale models.
- New isentropic analysis grid bundle
- Re-enabled custom colormaps
- GSD ensemble grid display added to OS X client
- Updated LDM to 6.12.14
- New bundles for ESTOFS and ETSS
    - Extra Tropical Storm Surge
    - Extra Tropical Combined Storm Surge and Tide
    - Ocean Surface Elevation

15.1.3, released March 21, 2016

- Full global topographic map (finally!)
- Maps database and shapefile updates from the newly-created awips2-static repository.
- Maps database RPM now updates itself correctly (in the past it would skip update/install if the database already exists, now it overwrites).
- Expanded grib1->grib2 variable mapping to support FNMOC grids.
- New WaveWatch/WW3 bundles:
    - Sig. Wave Height
    - Wind Wave Height
    - Swell Height & DIrection
    - Peak Wave Period
    - Sfc Wave Wind
    - Primary Wave Period
    - Secondary Wave Period
    - Sfc Vorticity and Wind
    - 4-panel displays of the above
- UKMET global grid compositing (replacing the previous N Hemisphere-only compositing)
- New LDM pattern actions available:
    - DGEX Alaska 12km
    - GFS Pacific 40 km
    - GFS Pacific 20 km
    - GFS CONUS 80 km
    - GFS CONUS 20 km
    - GFS Alaska 20 km
    - NAM CONUS 80 km
    - NAM CONUS 20 km
    - NAM Puerto Rico Grid 237
    - NAM Polar 90 km
    - AK NamDNG 5km
    - PR NamDNG 5km
    - Hawaii NamDNG 5km
    - AK NamDNG 3km
    - Gridded Excessive Precipitation
- New grids supported via changes/additions to the baseline:
    - HPCqpfNDFD
    - MOSGuideExtended
    - FNMOC NAVGEM
    - FNMOC NCODA
    - FNMOC FAROP
    - FNMOC COAMPS Regional Grids
    - MPE/QPE Puerto Rico
- New CAVE bundles:
    - High Frequency Radar

Ongoing Activities

Cloud Server Updates

- Now ingesting and serving 20+ gridded models due to some curious performance increases that happened on the Azure cloud.
    - DGEX
    - ECMWF
    - GFS 0.5 Global
    - GFS 40km
    - NAM 12km
    - NamDNG 2.5km
    - NAVGEM
    - NCWF
    - RAP 13km
    - RAP 40km
    - RTMA 2.5km
    - URMA 2.5km
    - UKMET Global
    - HRRR
    - Global Wave Watch III
    - ESTOFS
    - ETSS
    - FFG
    - fnmocWave
    - HFR
    - HPC / RFC / QPF / QPE grids

16.1.4 release is being prepared

- Requires clients and servers be updated together because 15.x and 16.x are not cross-compatible.
- More map projections.
- Himawari-8 support.
- Better support for native projection McIDAS files.
- New maps database and shapefile imports for StormSurgeWW and NHAdomain

GEMPAK

GEMPAK v7.2.3 was released in early 2016, incorporating NCEP table and map updates as well as expanding support for Python access to remote EDEX data. Originally the build strictly required the AWIPS II Python packages be installed, but it now builds against system Python and only requires the python-awips (DAF) be installed for GEMPAK programs to accepts prepended aliases for gridded models, such as A2GFS, A2NAM, A2HRRR, etc.

New Activities

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

- Investigate again the refactoring of pypies for h5py 2.6 compatibility.
- Collaborate with the Unidata Python team to integrate python-awips into workshop training materials and possibly into existing Unidata technologies.
- Move python-awips array manipulation to callable methods.

Relevant Metrics

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

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Both *AWIPS II and GEMPAK are freely available, and both incorporate LDM/IDD technology for efficiently accessing geoscience data. 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 II and GEMPAK are open-source, and while GEMPAK is now in maintenance mode, AWIPS II 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 II software package.

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

Status Report: Cloud Computing Activities

September-April 2016

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

Activities Since the Last Status Report

Docker

With the goal of better serving our core community and in fulfillment of objectives articulated in Unidata 2018: Transforming Geoscience through Innovative Data Services , Unidata is investigating how its technologies can best take advantage of cloud computing. To this end, we have been employing Docker container technology to streamline building, deploying, and running Unidata technology offerings in cloud-based resources. Specifically, we have created Docker images for the IDV, LDM, ADDE, RAMADDA, THREDDS, GEMPAK and Python and 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 Unidata's github repository.

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

Microsoft Azure for Research Grant and AMS "UniCloud: Docker Use at Unidata" Presentation

Our efforts in the Docker arena were presented at 2016 AMS annual meeting in New Orleans in a presentation entitled Unicloud: Docker Use at Unidata. In coordination with this presentation, we staged three "motherlode class" machines (in reference to our motherlode data server at Unidata) on our Microsoft Azure for Research resources; unidata-server, unidata-server-2, unidata-server-3. These servers provide data supplied by the LDM, and served by RAMADDA, TDS, and ADDE. They can be staged in minutes on cloud virtual machines with Docker and instructions for doing so can be found here.

Our Microsoft Azure for Research equipment grant will be ending mid-April 2016. We plan to respond to the April 15th Azure for Research RFP with several new proposals for Azure resources.

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. 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 very early stages of experimenting with Jetstream.

AWS Training/Technical Discussions at the University of Wyoming

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

Progress has been made on the following:

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

Azure for Application Streaming/Unidata Service Hosting

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

We have made available an EDEX Data Server in the Azure cloud (edex-cloud.unidata.ucar.edu), and have set up a similar server privately for Embry-Riddle Aeronautical University on an Amazon EC-2 instance. This Azure EDEX machine serves data to CAVE clients for Linux, Mac, and Windows, as well as Python scrip.
t and projects using the AWIPS II Python Data Access Framework (python-awips) and the latest GEMPAK build (which uses python-awips request Python data arrays objects and convert them into a renderable GEMPAK grid.

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 Dockerized version of the IDV, “CloudIDV”
- We have released a generic application-streaming container for use by our community with their own legacy software, “CloudStream”.
- We have deployed numerous services and instances to the Azure Cloud, mirroring our experiments with the Amazon cloud infrastructure.
- We have learned even more about Microsoft’s Azure cloud infrastructure.
- Submitted a talk to DockerCon 2016, being held this Summer, regarding the work Unidata has been doing with Docker.
- Staged three motherlode class machines on the Azure cloud with LDM, ADDE, RAMADDA, and THREDDS.

Ongoing Activities

We plan to continue the following activities:

- Use the LDM to move NEXRAD Level II data into AWS S3 buckets in real-time
- Develop enhanced procedures for recombining chunks of Level II data relayed in the IDD into full volume scans
- Develop TDS access to data stored in S3
- Maintain the TDS on AWS serving level II radar data
- Deploy iPython notebooks that provide access to NEXRAD Level II data stored in S3
- Deploy iPython notebooks that provide access to AWIPS II HDF5 data stored in the cloud.
- Expand the number of GEMPAK-supported data types requested from cloud-based EDEX servers.
- Continue experimenting with our Azure resources for running Unidata technology in the cloud and staging motherlode class machine.

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.
- Present a talk regarding our Docker work at DockerCon 2016 in June, if the proposal is accepted (announcements will go out at the end of March).
- Deploy a distributed EDEX installation in the cloud to ingest and serve the entirety of IDD gridded data sets.
- Develop AWIPS II EDEX access to S3 storage.
- Investigate the Jetstream cloud for running Unidata technology in the cloud.
- 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

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:

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

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

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

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

4. **Build, support, and advocate for the diverse geoscience community**

[Build a bigger community]

Prepared *April 2016*

Status Report: Community Services

October 2015 - March 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:

- Roland Stull Meteorology Text Available to Use at No Charge
- AWIPS II CAVE for Mac OS X
- NEXRAD Archive data available on Amazon S3
- Scott Jacobs Receives 2015 DeSouza Award
- CloudIDV: the Integrated Data Viewer in your web browser
- OGC Approves CF-netCDF Encoding Standard
- Unidata releases AWIPS II 15.1.2
- CloudStream — An Application Streaming Docker Framework
- Unidata Joins the Open Commons Consortium
- Welcome Communications Intern Larissa Gordon
- 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:

- Assisted the Users Committee in conducting the Unidata 2016 Community Survey (to be discussed separately at the Spring 2016 Users Committee meeting).
- A pilot project to assist community members with data management and document the resulting workflows for the DMRC was awarded supplemental funding by NSF. Three academic volunteers are participating (North Carolina State University, Millersville University, and the University of Wyoming). As a part of this project, community services intern Larissa Gordon has joined the UPC through summer 2016.

- Community engagement at recent professional society conferences (AGU, AMS, AGU Ocean Sciences)
- Engagement with CUAHSI to support the NFIE and WRF-Hydro at the NWC
- Continue to serve on the CUAHSI HIS standing committee
- Enhancing the Teaching Resource Network
- We continue to update Unidata's social media channels (Facebook, Twitter, Google+)
- We continue to publish short videos/screencasts on the Unidata YouTube channel.
- We continue to actively support the NCAR/SOARS program
- Actively participate in Super Science Saturday

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 and ESIP communities
- NAWIPS migration to AWIPS II, including the overall AWIPS II project
- Ongoing development of news articles for publication through News@Unidata
- Continue to support and contribute to governing committees
- Seminars
- Outreach
- Engagement with professional societies
- Support for cloud-related projects
- Further development of the Data Management Resource Center
- Further work on Agile Data Curation
- Site visits as the budget allows
- Engage other UCAR/NCAR divisions regarding Unidata software use i.e. CESM/IDV
- Continued additions to the Teaching Resource Network

New Activities

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

- Expanded emphasis on cloud-related activities
- Provide an initial online training resource focused on Python and Unidata services and tools
- Consult with NOAA NWS partners regarding the outcomes of the NWS SOO survey
- Release robust case studies for the DMRC with associated outreach through social media and the Unidata blog
- Participate in the Science of Team Science conference and identify possible collaborators

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

- Support Committees work towards producing the next Strategic Plan
- Feature the DMRC case studies at conferences
- Engage community to contribute curriculum and data to the Teaching Resource Network
- Continue to engage the hydrologic community regarding WRF-Hydro/IDV interactions and the National Water Center's efforts

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

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. The 2016 Unidata Users Committee community survey was recently completed; however, this is a self-selecting group of individuals that respond to a broad request for participants. Are there particular communities or groups that would benefit from individual outreach (e.g. either staff or committee member call for one-on-one discussions)?

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:

- 44,755 unique pageviews (39,360 in previous period)
- 7.9% of total unique pageviews to site (6.9% in previous period)

Top community pages

1. All blog pages
   33,825 unique pageviews (19.574 in previous period)

76% of total community pageviews (50% in previous period)
2. www.unidata.ucar.edu/community/
   4119 unique pageviews (3164 in previous period)
   9% of total community pageviews (8% in previous period)
3. www.unidata.ucar.edu/events/
   3333 unique pageviews (12010 in previous period)
   7% of total community pageviews (31% in previous period)
4. www.unidata.ucar.edu/about/
   2800 unique pageviews (3714 in previous period)
   6% of total community pageviews (9% in previous period)

Social media statistics, March 31, 2016

1. # of Twitter followers: 503 (up from 450)
2. # of Facebook followers: 490 (up from 428)

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.

Prepared *September 2015*

Status Report: Community Equipment Awards

Sponsored by the National Science Foundation

October 2015 - April *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 has been formed and will meet on April 12 at the Unidata Program Center. Applicants will be notified of award status by May 2016.

Areas for Committee Feedback

Relevant Metrics

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

Prepared March 2016

Status Report: Internet Data Distribution

September 2015 - March 2016

Mike Schmidt, Jeff Weber, Steve Emmerson, Tom Yoksas

Activities Since the Last Status Report

Internet Data Distribution (IDD)

0.25 degree GFS data was added to the CONDUIT data stream starting with the 12Z run on July, 28 . Monitoring has shown that peak CONDUIT data volumes increased from approx. 8 GB/hr to approx. 21 GB/hr for all forecast hours for the 0.25 degree GFS.

```
Data Volume Summary for emo.unidata.ucar.edu

Maximum hourly volume  38312.123 M bytes/hour
Average hourly volume  19769.266 M bytes/hour

Average products per hour     327728 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)            (M byte/hour)   number/hour
CONDUIT                  7878.758    [ 39.854%]    21708.893    89182.364
NGRID                    5206.945    [ 26.339%]     8596.371    37461.750
NEXRAD2                  2836.811    [ 14.350%]     4648.726    41083.250
NEXRAD3                  1578.781    [  7.986%]     1886.712    89978.795
FNMOC                    1274.134    [  6.445%]     3163.891     3624.205
HDS                       282.288    [  1.428%]      579.506    19087.614
NOTHER                    274.371    [  1.388%]      783.128     1360.091
NIMAGE                    155.168    [  0.785%]      251.517      181.523
GEM                        76.094    [  0.385%]      474.508      792.295
IDS|DDPLUS                 65.822    [  0.333%]       81.338    44297.227
EXP                        47.827    [  0.242%]       91.171      328.068
FNEXRAD                    45.000    [  0.228%]       83.920       43.614
UNIWISC                    44.757    [  0.226%]       90.713       21.386
LIGHTNING                   2.395    [  0.012%]        6.377      284.977
GPS                         0.113    [  0.001%]        0.997        1.045
```

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

Ongoing Activities

We plan to continue the following activities:

- Unidata continues to receive High Resolution Rapid Refresh (**HRRR**) grids (both 2D and 3D fields) in an LDM/IDD feed from NOAA/GSD and feed these products to a small number (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 currently receiving the NOAA/GSD data are amenable.

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

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

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

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

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

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

- The UPC continues to relay FNMOC and CMC data model output directly to the community. FNMOC provides the COAMPS and NAVGEM model output and the CMC provides the GEM model output. Unidata has provided access to these data for the past 8 years, but on a "point-to-point" basis. GEM model output was converted from GRIB1 to GRIB2 in January, 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 datastreams (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. Attempts to add a Northrop Grumman office in Northern Virginia was not successful due to their having very limited network bandwidth. We are still seeking a fourth ingest site to increase robustness of the IDD.

- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. A new LDM release, v6.13, will be released *real soon now*.

- Raytheon continues to submit modifications to the LDM for enhanced AWIPS-II functionality.

Relevant Metrics

- Approximately **565** machines at **253** sites are running LDM-6 **and** reporting real time statistics to the UPC. Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

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

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

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

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

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

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

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.

Prepared *March, 2016*

Status Report: IDV with RAMADDA

September-April 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 4.6.2 (June, 2015). Its main feature is a rewrite of GRIB handling. This version of netCDF-Java includes changes: AWIPSsat projection, WRF lat/lon coordinates with time dependence, updated grib1 table for NLDAS, handling of scalar runtime coordinate for GRIB collections. See the netCDF-Java Library for a more details on these changes.

__ISL Changes__

- Added ability to export XIDV to ZIDV.
- Added ability to choose timezones when displaying times.

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 at workaround for Java 3D grey window/panel problem
- Added calibration to local McIDAS area files, currently only for the GVAR area file.

__Other Changes__

- Color palette improvements for movie capture.
- Various improvements for textual dialogs in the IDV. (Pull requests 99, 105, 106.)
- The webcam database was updated and we are now serving these images from the Azure cloud.

IDV CESM Demonstration

In March 2016, a well received IDV demonstration was presented by Yuan to the CESM (Community Earth System Model) group. **They are requested by their funding agency** to use IDV for their model output visualization.

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 January 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 Turbulence Demonstration

With Ryan's Python script to process the fortran binary array, the IDV is able to demonstrate the fascinating turbulence vertex structure.

IDV Release

The last IDV release was 5.2 in August of 2015.

IDV Proposal Work

__EarthCube__

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"

__NASA ROSES__

In July of 2015, with our University of Wisconsin, Space Science and Engineering Center partners, we submitted a proposal to NASA entitled, "Interactive Algorithm Development and Product Validation through Innovative Data Access and Visualization Methods". We are still awaiting a response from that proposal.

IDV Publication Highlights

A Google Scholar Search reveals a number of publications that cite use of the IDV.

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. This work was presented at 2016 AMS annual meeting in New Orleans in a presentation entitled Unicloud: Docker Use at Unidata. We staged three demonstration servers on the Azure cloud. In this example, RAMADDA is serving IDD data supplied via the LDM.

IDV and RAMADDA Training, Conference Attendance and Presence

__2016 American Meteorological Conference (AMS) Annual Meeting__

- Visualize the high resolution model output with 3-D Adaptive Resolution feature in the IDV
- Meteorological Applications in a Collaborative Classroom and Stereo Visualization Laboratory
- UniCloud: Docker Use at Unidata
- RAMADDA on the Azure Cloud demonstration server
- IDV presence at the Unidata booth and student career fair

__2015 American Geophysical Union (AGU) Annual Meeting__

- Integrating Satellite, Radar and Surface Observation with Time and Space Matching
- Bringing the Unidata IDV to the Cloud
- IDV presence at UCP booth

__IDV Instructional Videos__

Produced a new IDV training video entitled Running the IDV on the XSEDE Jetstream Cloud Demo.

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 ~100 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 97 “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 over 8,000 times compared with 6,115 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. 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

October 2015 - March 2016

Tom Yoksas, Yuan Ho

Activities Since the Last Status Report

No new activities

New/Planned Activities

Training Workshop at the Chinese Meteorological Agency offices in Beijing

Unidata will conduct a multi-day training event at the Chinese Meteorological Agency offices in Beijing, China in mid-September.

The training will focus on the use of the LDM to distribute meteorological data, the IDV to visualize data, and use of GOES imagery.

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 *March, 2016*

Status Report: LDM

October 2015 - March 2016

Steve Emmerson, Tom Yoksas, Mike Schmidt

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:

- LDM logging was rewritten to eliminate dependence on the system logging daemon because:
  - It was the major impediment to correctly using the LDM;
  - It was the major impediment to putting the LDM in a Docker container;
  - Only 7 "local" syslog(8) facilities are guaranteed and NOAAPort ingestion could use more; and
  - Some system logging daemons would drop LDM log messages depending on how heavily they were loaded.
- Missing parameter entries were added to the GEMPAK tables for NOAAPort ingestion.
- The possibility of `pqact(1)` not processing the last successfully-received data-product due to a restart of the LDM was eliminated.
- Access to the product-queue was made thread-safe.
- Raytheon's changes to NOAAPort ingestion to properly handle GRIB-2 messages that mix compressed and uncompressed frames were incorporated.
- Version 6.13.0 was released.

Dependencies, challenges, problems, and risks include:

- Dealing with missing GEMPAK table entries is a pain.
- Very low bus factor (if Emmerson gets hit by a bus, the LDM's in trouble).

Multicast LDM (alias LDM-7)

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

Progress has been made on the following:

- An oral presentation on LDM-7 was made at last year's Fall AGU meeting.
- Testing of LDM-7 over static VPN circuits on Internet2 has begun. Participants include University of Virginia, Rutgers University, Indiana University, and University of Wisconsin.

Dependencies, challenges, problems, and risks include:

- Time.
- Very low bus factor.

Ongoing Activities

We plan to continue the following activities:

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

New Activities

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

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

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

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

Relevant Metrics

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

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.

Prepared *March 2016*

Status Report: McIDAS

October 2015 - March 2016

Tom Yoksas

Activities Since the Last Status Report

Aside from routine updates/bugfixes to existing code and tables, the main thrust of development is to add indexing to ADDE datasets to speed access into large datasets.

Prior Activities

- Unidata McIDAS version 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 Inquiry Metrics

ldm-mcidas Decoders Activities

Development

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

Geostationary Satellite Data Ingest

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

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

Planned Activities

Ongoing Activities

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

New/future Activities

Install a GOES-R downlink, processing and data serving capability at the NCAR Mesa Lab using the easternmost of the old USAN satellite pads. NOAA GOES-R office funding for this activity should become available *real soon now.* Launch of GOES-R is currently planned for October, 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**
   *McIDAS remains **the** application of choice for the satellite meteorology community. The Advanced Data Distribution Environment (ADDE) component of McIDAS was the first application offered by Unidata to provide remote, programmatic access to a wide variety of data that is important to the atmospheric science community.*

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

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

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

Prepared *March, 2016*

Status Report: netCDF

September 2015 - April 2016

Ward Fisher, Dennis Heimbigner, Russ Rew

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 17 open issues for netCDF-C, 9 open issues for netCDF-Fortran, and 2 open issues for netCDF-C++. The netCDF Java interface is maintained by the Unidata CDM/TDS group (which also uses Jira and GitHub), and we collaborate with external developers to maintain the netCDF Python interface.

In the netCDF group, progress has been made in the following areas since the last status report:

- We have included support for 64-bit-everything netCDF format from the parallel netCDF project at Argonne and Northwestern into a branch of the netCDF project.
- Extend 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 upcoming hdf5 1.10 version has been advertised as providing a different file format and suite of capabilities, compared to the hdf5 1.8 version. The scope of the changes are unknown, but the potential impact may be either very small, or very large. This is yet unknown.

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.
- Release the next versions of netCDF-C, netCDF-Fortran, netCDF-C++.
- Modernize the netCDF documentation to provide easy access to documentation for older versions of netCDF.

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

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

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

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

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

Relevant Metrics

There are currently about 140,500 lines of code in the netCDF C library source.
The Coverity estimate for defect density (the number of defects per thousand lines of code) in the netCDF C library source has been reduced slightly from **0.34** six months ago to **0.32** today. According to Coverity static analysis of over 250 million lines of open source projects that use their analysis tools, the average defect density with 100,000 to 500,000 lines of code is **0.50**.

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

Currently, Google hits, for comparison, are:

- **535,000** for netCDF-3
- **517,000** for netCDF-4
- **381,000** for HDF5
- **69,500** for GRIB2

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

- **237** for netCDF-3
- **389** for netCDF-4
- **6,520** for HDF5
- **545** 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

September 2015 - April 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 telcons
- Arranged for Unidata to host ODIP Workshop in May 2016.

Dependencies, challenges, problems, and risks include:

- There is a question of what resources should be committed to these projects beyond the end of the 2016 calendar year.

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:

- Held kick off teleconference to outline steps for getting a “motherlode clone” running on the GMU EarthCube hardware.
- Maintained periodic contact with Liping Di, the GMU PI on the project.

Dependencies, challenges, problems, and risks include:

- The project got off to a slow start but the PI, Liping Di of GMU, anticipates that a one year no cost extension will make it possible to achieve the proposed objectives.

Python Workshop Notebooks on alternative platforms

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

cloud.

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.

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

We are requesting your feedback on the following topics:

Should we pursue an OGC standard for conventions for including ISO conforming metadata in netCDF files?
Regarding a "short cut" OGC standard for GRIB, is it necessary to spell out systematic online access to external tables and to spell out explicit data model mappings?
How important is it to continue these outreach efforts beyond the end of the 2016 calendar year. Are some of these projects higher priority than others?

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.

Prepared *April 2015*

Status Report: Python

September 2015 - March 2016

Ryan May, Sean Arms, Julien Chastang, Ward Fisher

Activities Since the Last Status Report

Siphon

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

Progress has been made on the following:

- Expanded support for netCDF extended model through CDMRemote
- Support for CDMRemote v2 from TDS 5
- Community contribution of fixing Windows-only bug in NCSS handler
- Improved automated code quality checking with Codacy and QuantifiedCode
- Improved infrastructure of testing (including porting from nose to py.test)
- Changes to ensure automated documentation builds do not silently fail
- Addition of automated testing on Windows using AppVeyor

MetPy

Feedback on MetPy continues to be positive. We have been publicizing the project in a variety of ways:

- Ryan May presented on MetPy at the Python Symposium at the 2016 AMS Annual Meeting.
- The Unidata Developers' blog had a two part post using Siphon and MetPy to plot GINI satellite data
- The MetPy twitter account has reached 41 followers
- Ryan May and Sean Arms (with input from Kevin Goebbert) will present a tutorial on MetPy and Siphon at the Software Engineering Assembly conference.

There have been some external contributions from the community, including a Pull Request adding a function to reduce sounding data to a desired set of data. Ideas for further development are captured in the GitHub issue tracker; current plans are to focus on enhancing the abilities to read point data (to facilitate generating NetCDF data for hosting on THREDDS) and to objectively analyze such data.

Progress has been made on the following:

- Continued investment in automating build, testing, documentation, and release infrastructure using Travis, Codacy, QuantifiedCode, and Read The Docs
- Added image-based tests for automated testing of plotting functionality

- Added support for reading GINI satellite imagery, including a netcdf-like interface
- Implemented station plots including weather symbols
- Improvements to Skew-T plots and added hodographs
- Community awareness and involvement progressing well one year into the project

Dependencies, challenges, problems, and risks include:

- Due to small dedicated staff time, progress adding to MetPy has been slower than desired.

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.

Progress has been made on the following:

- Fixed wind barbs for the upcoming 2.0 release of matplotlib
- 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

Unidata Python Workshop

We continue to improve and expand this popular workshop, now three days long, with new material, 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.

Python Online Training Effort

Unidata obtained supplemental funds from NSF as part of our five-year award to start an online Python training effort specifically focused on serving the geoscience community. The Unidata Python group has outlined sections for introductory material, and is currently working towards having a draft resource in the early summer.

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

New Activities

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

- Using supplemental funds from NSF, develop asynchronous training materials for Python in meteorology. We are investigating the use of a cloud server hosting executable Jupyter Notebooks (based on our training workshop) as the core of the training materials, using either the tmpnb or jupyterhub packages from Project Jupyter.
- CIMSS/SSEC has asked Unidata to present our Python Training Workshop in a pair of two-day workshops in Madison, WI in June.
- Investigate the use of having THREDDS communicate directly with a Jupyter/IPython kernel for server-side processing functionality

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

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. What are the biggest obstacles that you see to the use of Python with other Unidata technologies, or for use in meteorology in general?
2. How valuable do find an effort like MetPy to the Python meteorology community? Are there additional barriers we could remove through this project? Are there other efforts over which this should take priority?
3. Have you seen the two-part blog post using MetPy and Siphon? How was the tone for a "notebook of the week"-type series?
4. 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
- 357 downloads/month from the Python package index
- Watchers: 10
- Since 1 October 2015:
  - Active Issues: 31 (19 created, 12 closed)
  - Active PRs: 23 (21 created, 22 closed)
  - External Issue Activity: 1 opened, 7 comments
  - External PR Activity: 0 opened, 4 comments
  - Unique external contributors: 5
  - Stars: 6 (20 total)
  - Commits: 138
- Since 1 April 2015
  - Active Issues: 55 (55 created, 31 closed)
  - Active PRs: 41 (41 created, 40 closed)
  - External Issue Activity: 6 opened, 16 comments
  - External PR Activity: 4 opened, 8 comments
  - Unique external contributors: 11
  - Stars: 14 (20 total)
  - Commits: 250

MetPy:

- 93% test coverage
- 798 downloads/month from the Python package index
- Watchers: 11
- Since 1 October 2015
  - Active Issues: 53 (20 created, 19 closed)
  - Active PRs: 29 (28 created, 27 closed)
  - External Issue Activity: 3 opened, 11 comments
  - External PR Activity: 3 opened, 1 comments
  - Unique external contributors: 8
  - Stars: 14 (58 total)
  - Commits: 153
- Since 1 April 2015
  - Active Issues: 77 (75 created, 41 closed)
  - Active PRs: 58 (58 created, 56 closed)
  - External Issue Activity: 11 opened, 35 comments
  - External PR Activity: 6 opened, 2 comments
  - Unique external contributors: 15
  - Stars: 53 (58 total)
  - Commits: 478

Unidata Python Workshop (from github):

- Watchers: 28
- Stars: 24

- Forks: 49
- Issues: 45 (40 created, 5 closed)
- Pull Requests: 19 (0 open, 19 closed)
- Master Branch Commits: 545

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 also fostering some community development of meteorology-specific tools under the MetPy grassroots project.

Status Report: Support

October 2015 - March 2016

Tom Yoksas, Jennifer Oxelson, UPC Staff

Activities Since the Last Status Report

Training

- The UPC hosted its annual training workshop series from July 20 - August 5, 2015 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 48500 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system.

Support by Category

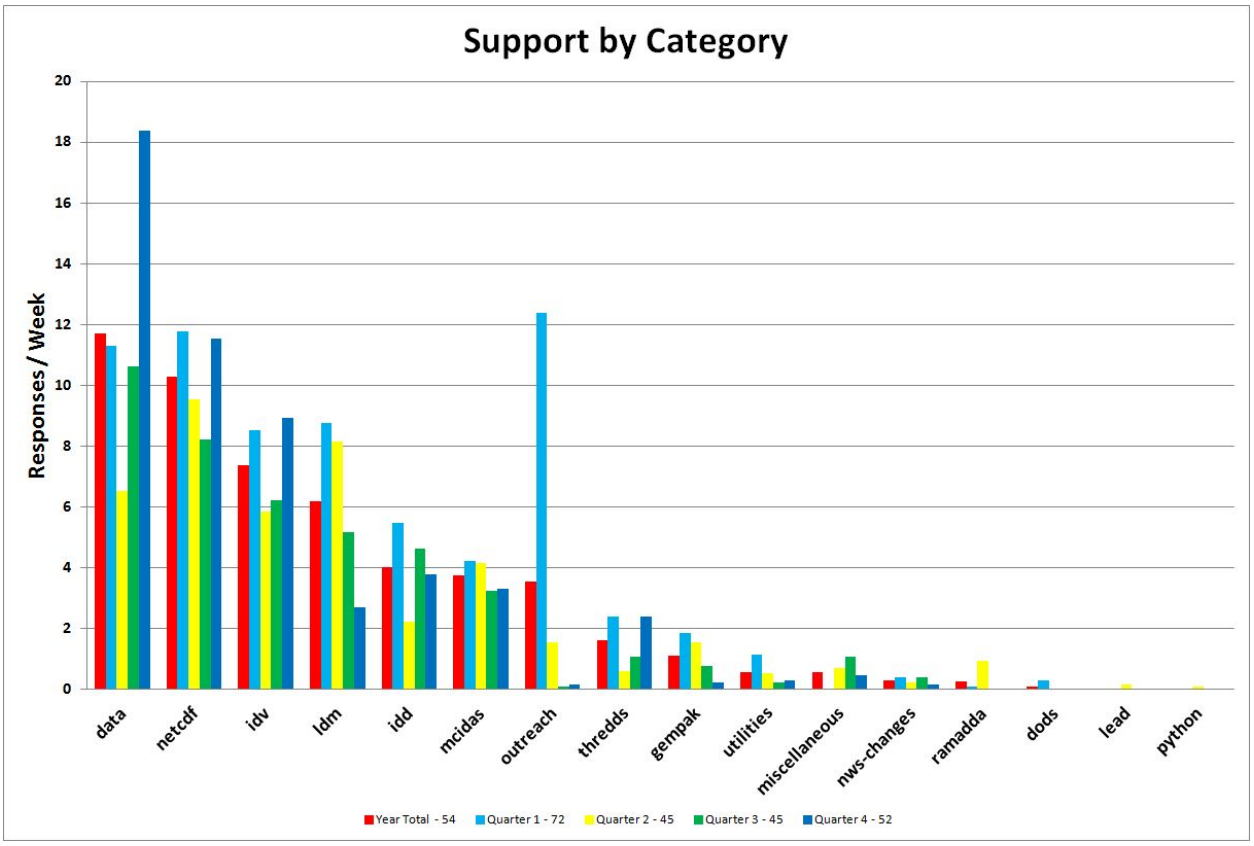

Fig. 1: Above are histograms that portray the number of Unidata email responses for categories of support for a one year period ending March 22, 2016. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period is depicted 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, fxlinux, misc, license, network, notrack, platforms, wxp wxp-lis
netcdf	data-models, libcf, ncml, netcdf, netcdf-miss, netcdfgroup-list, netcdf-hdf, netcdf-hdf-list, netcdf-java, netcdf-perl
nws-changes	nws-changes
outreach	agu-ty, announce, argentina-ty, barbados-ty, brazil-ty, cathalac-ty, chile-ty, costarica-ty, mexico-ty, support-ty, cbmet-ty, community-list, eletter, egrants, eumetsat-ty, external, iai-ty, international-ty, joss-ty, k12-list, korea-ty, meteoforum-ty, unidata, workshop
python	python
ramadda	ramadda
thredds	java-dev, java-dev-list, thredds
utilities	decoders, ldm-mcidas, udunits

Support by Topic

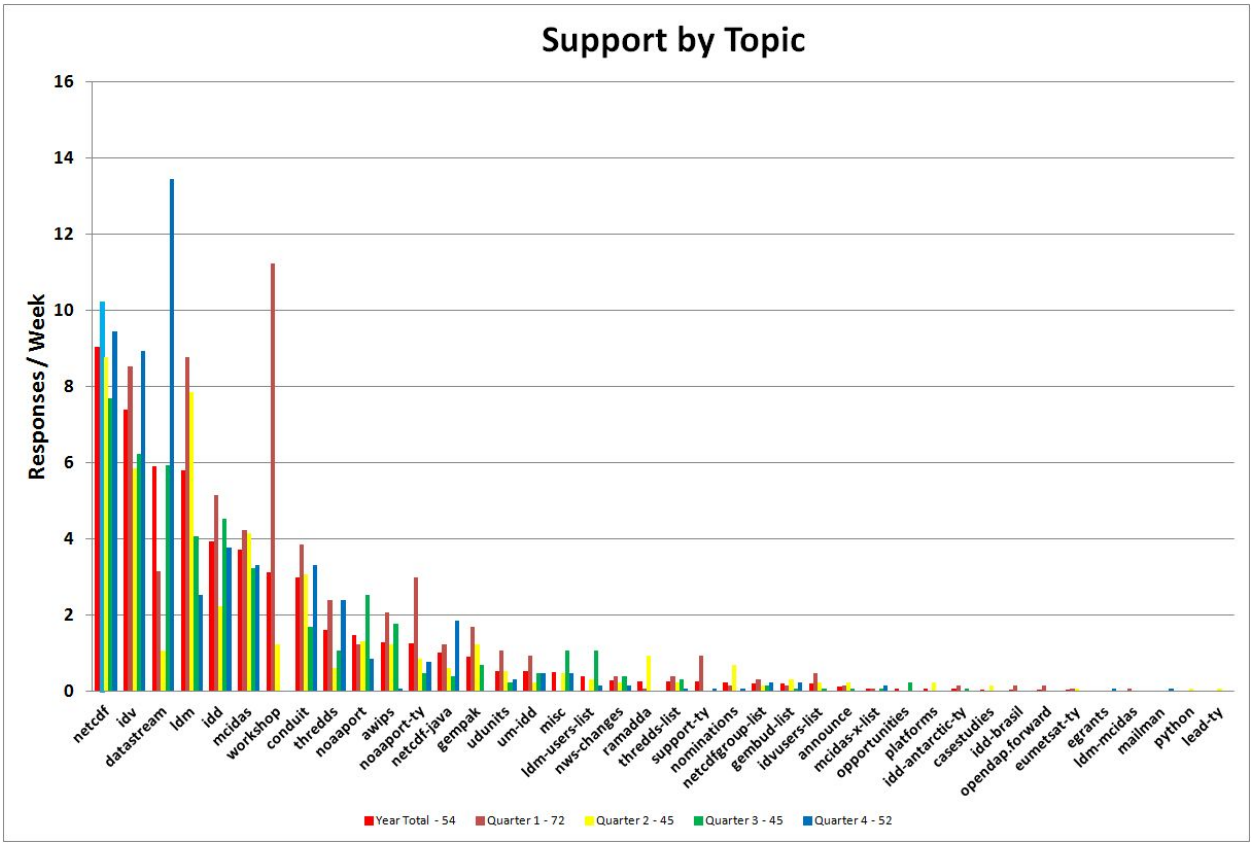

Fig. 2: Above are histograms that portray the number of Unidata email responses for individual topics of support for a one year period ending March 22, 2016. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period 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 netcdf-java continues to grow steadily.
- Support for the legacy visualization packages GEMPAK and McIDAS continues to be substantial. Support for AWIPS-II has been increasing steadily and now exceeds that for GEMPAK.

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

Notes

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

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

Additional User Support Inquiry Metrics

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 *March, 2016*

Status Report: THREDDS

September 2015 - April 2016

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

Activities Since the Last Status Report

The THREDDS Project

The THREDDS Project encompases four projects: netCDF-Java, the THREDDS Data Server (TDS), Rosetta, and Siphon (the Unidata Python client to interact with a TDS). For specific information on Siphon, please see the Python Status Report.

Released netCDF-Java / TDS version 4.6.5 (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

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
- Unit and Integration tests are passing

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:

- The discontinuation of the ACADIS project presents funding challenges for work on Rosetta. External funding is being pursued and collaboration opportunities are appreciated.

Ongoing Activities

We plan to continue the following activities:

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

The following active proposals directly involve THREDDS work:

- New EarthCube award: "Advancing netCDF-CF for the Geosciences". This two-year, Unidata lead project will work to extend netCDF-CF conventions in ways that will broaden the range of earth science domains whose data can be represented.
- Beginning the second year of NASA ROSES ACCESS award: "High Performance Multidisciplinary Open Standard Data Services to Serve Terrestrial Environmental Modeling" with USGS CIDA.
- Three EarthCube awards are finishing up on a no-cost extension:
  1) EarthCube Building Blocks award: "Integrating Discrete and Continuous Data" with Univ of Texas, Austin and others.
  2) EarthCube Building Blocks award: "Specifying and Implementing ODSIP, A Data-Service Invocation Protocol" with OPeNDAP, Inc.
  3) EarthCube Building Blocks award: "Deploying Web Services Across Multiple Science Domains" with IRIS, UNAVCO, and others. Period of performances: Oct 2013

- Sept 2015.
- Still awaiting decision on NASA ROSES ACCESS proposal: "Leveraging available Technologies for Improved interoperability and visualization of Remote Sensing and in-situ Oceanographic Data at the PO.DAAC" with JPL/PO.DAAC. [Rosetta]
- Submitted new DIBBS Early Implementation proposal with Texas Tech: “CIF21 DIBBs: EI: Enhance Unidata Data Facility With a Scalable Lightning Data Infrastructure and Geospatial Capabilities". [netCDF-Java/TDS/Siphon]

New Activities

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

- Releasing a beta of TDS 5.0
- Making public a TDS 5.0 Test Server
- 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
  - 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
- Move to a fully online based training tutorial, reserving in-person, annual training for advanced topics

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.
