Users Committee Fall 2018 Agenda

Unidata Program Center, Boulder CO

Monday, 10 September 2018

8:30 - 9:00 Doors Open

9:00 - 9:20 Welcome and Administrative Items (continental breakfast available)

- New Members
- Date for Spring 2019 meeting
- Review of Action Items from Spring 2018 meeting

9:20 - 10:20 Director's Report & Questions - Mohan Ramamurthy

10:20 - 10:30 Break

10:30 - 11:00 GOES 16/17 data with GEMPAK - Michael James

11:00 - 12:00 Discuss Unidata 2024 NSF core funding proposal

12:00 - 1:15 Lunch

1:15 - 1:45 Pangeo Update - Rich Signell, Ryan May, Ward Fisher

1:45 - 2:30 Status Reports - (Users Committee should review status reports prior to the meeting) - staff will be available for questions

2:30 - 2:45 Break

2:45 - 3:45 NOAA - Carissa Klemmer

3:45 - 4:00 Transition to Auditorium

4:00 - 5:00 Russell L. DeSouza Award and seminar

5:00 Adjourn

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

at Zolo Southwestern Grill, 2525 Arapahoe Avenue, Boulder, CO (map)

Tuesday, 11 September 2018

8:30 - 9:00 Doors Open

9:00 - 9:30 Around-the-table reports from members (continental breakfast available)
(Things you are currently working on related to Unidata projects/computing in general)

9:30 - 10:00 National Academy Report (& FAIR 1 and 2) and what it means for Universities - Group Discussion

10:00 - 10:30 Post-Triennial Lessons Learned - Josh Young & Triennial Co-Chairs

10:30 - 10:45 Break

10:45 - 11:45 Blue Sky Session

(Things you are dreaming that could be reality, but needs work to achieve that Unidata could potentially help with)

11:45 - 12:30 AOB/Review of Actions/Box Lunch

Status Report: AWIPS and GEMPAK

April 2018 - September 2018

Michael James

Activities Since the Last Status Report

AWIPS

Unidata's Jetstream EDEX server continues to serve real-time AWIPS data to CAVE clients and through the python-awips data access framework (API). By offloading the processing of one very large data feed (NEXRAD3) to a separate EDEX Ingest Node, the current implementation of **edex-cloud** is now cabable of processing and serving more data than ever before. The distributed architectural concepts of AWIPS allow us to scale EDEX in the cloud to account for the size of incoming data feeds.

AWIPS 17.1.1-6 release June 2018

- This is a bug fix release primarily for the **WarnGen** application. For practice sessions, WarnGen no longer uses JMS to send or verify messages between client and server, instead sending the generated message directly to a TextWS window for editing, keeping everything client-side (see #217)
- **postgis** update to 2.4.4 to fix segfault signal 11 when WarnGen requests EDEX PostgreSQL to create coverage area geometries (see #183)
- WarnGen-dependent libraries **geos**, **gdal**, **proj,** and **shapely** updated to latest versions.
- **Postgresql** and **psql** updated to 9.5.13

CAVE for Windows

We now have a stable CAVE development environment for Windows, with Java 1.8, Eclipse 4.6.1, Python 2.7. With a free educational use license provided by the Advanced Windows Installer developers, CAVE can be delivered in a custom MSI installer with Unidata branding - with this, a new CAVE client is available for Windows (currently in beta but available at http://www.unidata.ucar.edu/downloads/awips2/awips-cave-18.1.1.msi).

Build and Source Code

We are now managing AWIPS source code across **19** different Git repositories for the Unidata build, including new repos **awips2-hazards**, **awips2-swpc**, **awips2-ohd**, and **awips2-nativelib**.

For faster development and testing cycles, I have created a new Docker image **awips-devel** as

a base for the **awips-ade** build container.

Python for AWIPS

We continue to update the Python Data Access Framework (**python-awips**) package with the latest changes from the AWIPS baseline. This package is used in both AWIPS and GEMPAK for remote retrieval of AWIPS data (grids, geometries, and imagery), and is showcased in a set of Jupyter Notebooks available at https://nbviewer.jupyter.org/github/Unidata/python-awips/tree/master/examples/notebooks

The next **python-awips** release will be 18.1.1 to correspond with the upcoming release of AWIPS EDEX and CAVE (new naming convention from previous 0.9.11, 0.9.12, etc.)

NWS changes to python-awips which are being incorporated into the next release include:

- New class **GetGridLatLonRequest** for lat/lon requests to be made separately from gridded data requests, allowing for "lazy loading" of grids where the library waits until the lat/lon for a grid is needed before loading. Examples of when this is True are when the lat/lon information is not used, or when it is used if certain conditions are met. It could be set to False if it is guaranteed that all lat/lon information is needed.
- New setters and getters for geospatial envelope and coordinate reference system (WKT) in **GetGridDataResponse.py**.New **DataQueue** convenience class for using the notifications feature. This is a collection that, once connected to EDEX by calling start(), fills with data as notifications come in. Runs on a separate thread to allow non-blocking data retrieval.
- New data notification class **PyGridNotification** which improves performance over requesting by datatime since it requests only the parameter that the notification was received for (instead of this and all previous parameters for the same forecast hour). However, this still fails for derived requests.
- **Apache Thrift** updated to version 0.10.0.

GEMPAK/NAWIPS

GEMPAK release 7.4.2 April 2018

https://github.com/Unidata/gempak/releases/tag/7.4.2

- The program **gdpvsf** was added back into this release after a fix to address the maximum number of grids allows, which was causing a memory reallocation error.
- Added polygon-based **Snow Squall Warnings Product** (nwx, nmap2, gpmap) from NCEP NAWIPS.
- General table and map updates from NCEP NAWIPS 7.4.2.

Activities Ongoing/In-Progress

AWIPS

Archive Case Study Efforts

I have begun adding table and decoder support for older data sets, models, satellite sources, for purposes of representing COMET case study data sets in EDEX and CAVE. Efforts have been made to deploy a standalone AWIPS Archive EDEX server in the cloud to contain notable case study events, currently in the testing and development stage.

EDEX and WRF

Using the Jetstream cloud, I have begun work on an instance which uses the UEMS (WRF) to generate custom forecast grid domains with direct EDEX ingestion and processing. The hope is that the UEMS can be incorporated into EDEX as its own installable el6 and el7 RPM so that any EDEX server can turn-key generate and ingest its own WRF domains.

Unidata AWIPS release 18.1.1 (September 2018)

- New colorized plugin for **International Sigmets**.
- New roles/permissions framework (this will require we all update to 18.1.1 at the same time or else supper the Vizalert red error messages)
- Database reconfiguration (moved from **/awips2/data** to **/awips2/database/data**, now with new postgresql owner role 'awipsadmin').
- Fix for CAVE crash when starting **GFE** for non-activated site.
- Updates to allow dynamic contributions to the **Volume Browser.**
- New viz method **IGLMesh** for easier OpenGL extensibility.
- Ability to handle derived parameter cubes at a single point.
- Search By Name has been added to the **Localization Perspective**.
- Fix for decoding PNG-compressed UNIWISC McIDAS AREA files.

AWIPS Development Efforts

- I'm currently evaluation new **AWIPS Hazards** plugins for decision assistance (product generation tools and hydro configurations integrated with WarnGen and the Hydro perspective).
- Unidata now has the source code to build 64-bit **Hydro Applications** (these Hydroapps were supplied as 32-bit binaries in the Vlab source code repositories, which were disabled in previous Unidata AWIPS releases).
- **OHD** and **Hydro Applications** may be made available as optional EDEX add-ons in the awips2-edex-dat and awips2-hydroapps RPMs.
- Also currently building and testing **SWPC plugins** for Geomagnetic Data, Solar Imagery, Generic High Cadence Data, and Time Series displays).

- Testing **OGC** and **Data Delivery** plugins to be made available after the initial 18.1.1 release as an optional add-on to an EDEX installation (two additional RPMs).

GEMPAK

GEMPAK release 7.5.1 September 2018

- Bug fix for GOES ABI imagery files in McIDAS AREA format (UNIWISC)
- Added NCEP support to handle Himawari data in netCDF4 format
- New lutfils for GOES 16 IR and WV channels (https://www.unidata.ucar.edu/staff/mjames/gempak/goes16/)
- Corrected imgtyp.tbl for Himawari-8 image type numbers
- Updated Eastern North Pacific and Atlantic Basic hurricane forecast cone radii for 2018
- Corrected a bug in 7.4.2 snow squall warnings which caused dcwarn to generate false FFW warnings
- Various table and map updates from NAWIPS 7.5.1
- Updated zlib 1.2.6 -> 1.2.8
- Updated HDF5 1.8.0 -> 1.8.15
- Updated netCDF 4.1.3 -> 4.3.3.1

New/Future Activities

- Test AWS API for public data set retrieval via EDEX.
- Expand AWIPS geospatial city table to include major world cities.
- Investigate how the Weather Event Simulator 2 Bridge (WES-2 Bridge) can be (and if it should be) merged into the Unidata AWIPS baseline.
- Migrate more NCEP Viz plugins from the NCP to D2D.
- Investigate creating an AppImage execulable for CAVE for wider Linux distro support.
- Add networking to the CAVE Archive Case Study creation tool.
- Incorporate USER-level bundles and modifications in the D2D data selection menus

Metrics

AWIPS downloads
GEMPAK downloads
Python-AWIPS downloads

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Both AWIPS and GEMPAK are freely available, and both incorporate LDM/IDD technology for accessing geoscience data. The cloud-based EDEX data server continues to see widespread use and growing adoption. More and more datasets continue to be added to the server as Unidata deploys more decode/ingest nodes.

2. **Develop and provide open-source tools for effective use of geoscience data**
   Both AWIPS and GEMPAK are open-source, and while GEMPAK is now in maintenance mode, AWIPS is continuously being developed.

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

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

Prepared *August 2018*

Status Report: *Cloud Computing Activities*

April 2018 - September 2018

Julien Chastang, Ward Fisher, Michael James, Ryan May, Jen Oxelson, Mohan Ramamurthy, Jeff Weber, Tom Yoksas

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. What clouds are our community using, either commercial (e.g., Amazon) or non-for-profit (e.g., NSF XSEDE Jetstream)?
2. What new cloud technologies are our community using and investigating on their own initiative?
3. Who would like to volunteer to beta test CloudIDV?
4. Have you had a chance to try the Unidata JupyterHub server?
5. Do you need a Unidata hosted JupyterHub on Jetstream for your classroom or workshop use?

Activities Since the Last Status Report

Successfully Awarded New Research Allocation on NSF Jetstream Cloud

Unidata successfully requested and obtained a renewal to our Jetstream allocation worth $530,000 (compared with $425,000 from the previous award):

- Awarded Resources: IU/TACC (Jetstream): 3,560,940.0 SUs
- IU/TACC Storage (Jetstream Storage): 40,000.0 GB
- TACC Data Analytics System (Wrangler): 8,640.0 Node Hours
- TACC Long-term Storage (Wrangler Storage): 40,000.0 GB
- The Science Gateways Community Institute (SGCI) support
- XSEDE Extended Collaborative Support Services (ECSS)

JupyterHub Servers on Jetstream Requests

At the ESIP 2018 Summer Meeting in Tucson, AZ, Unidata organized a session entitled Science Gateways in the Cloud, a Platform for Providing Modern Scientific Workflows for Reproducible Research and Collaboration. Julien Chastang gave a presentation on a Unidata Science Gateway JupyterHub. Since that presentation, we had four requests to host JupyterHub servers on Jetstream for classroom and conference use of which we have fulfilled three:

- Tom Narock | Notre Dame of Maryland University

- Keith Maull (UCP) | Southern Arkansas University
- JupyterHub for 2018 IS-GEO Workshop

We also turned down hosting ESIPHub server on Unidata's Jetstream allocation because we felt it was out of scope.

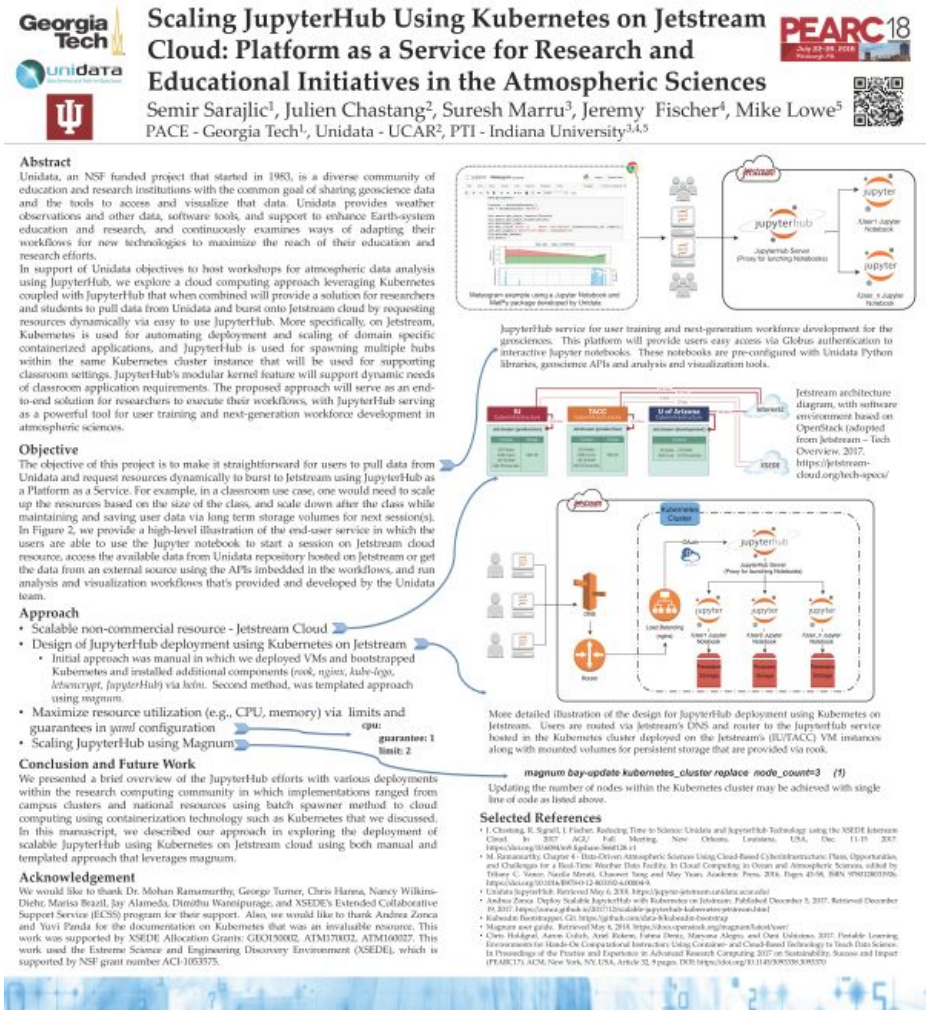
Dependencies, challenges, problems, and risks include:

- For classroom and workshops with < 10 users, setting up these JupyterHub servers is relatively straightforward. However, scaling these servers to accomodate a large number of users remains an active area of investigation. (We will be presenting on this topic at the 2019 AMS Annual Meeting.)

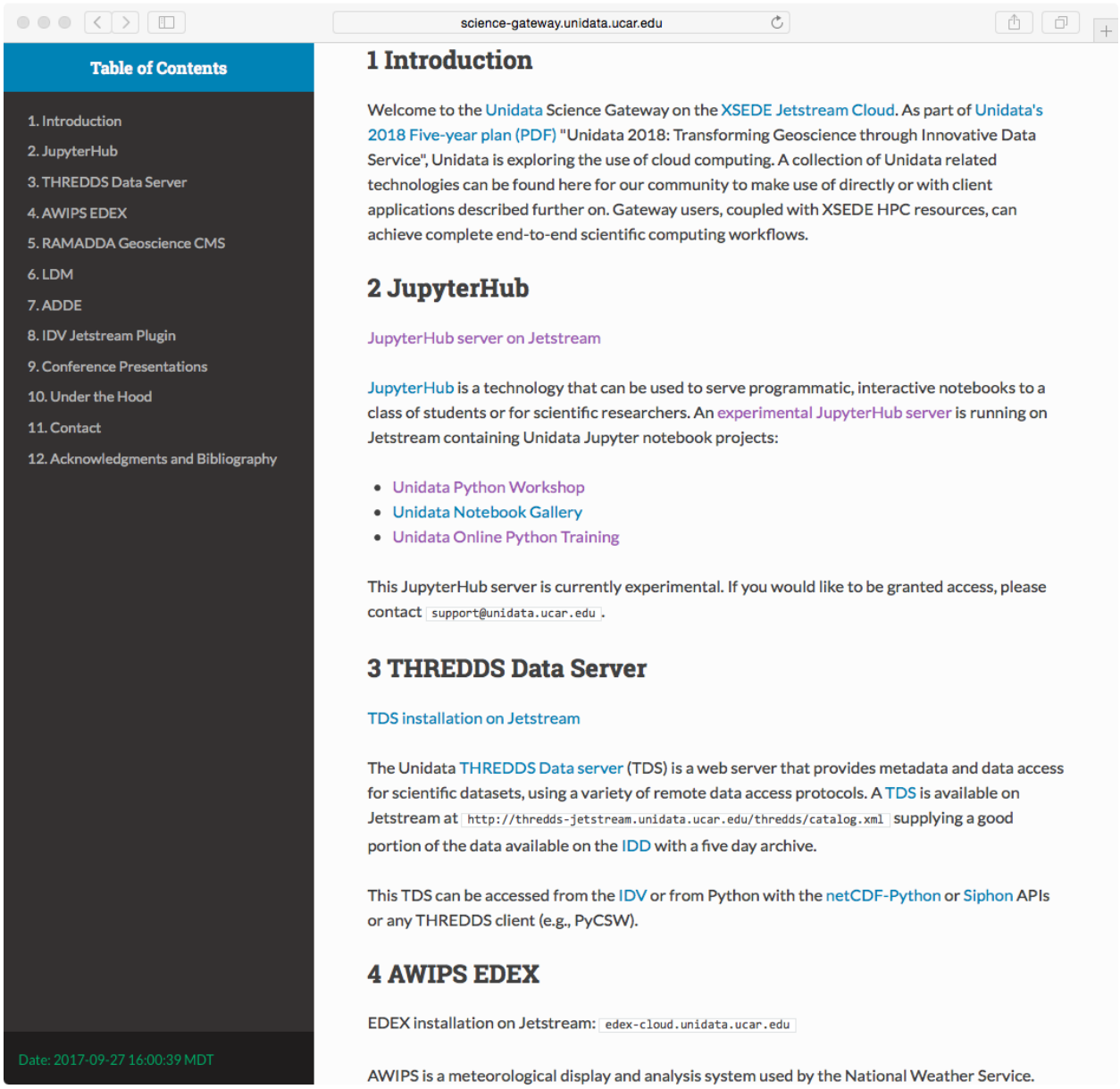
Unidata XSEDE ECSS Collaboration and Best Poster PEARC18

Julien Chastang and collaborators recently won the best poster award at the Practice & Experience in Advanced Research Computing (PEARC18) conference for their poster entitled "Scaling JupyterHub Using Kubernetes on Jetstream Cloud: Platform as a Service for Research and Educational Initiatives." The award is a milestone in a collaboration of several years between Unidata, the eXtreme Science and Engineering Discovery Environment (XSEDE) Extended Collaborative Support Services (ECSS) team and Indiana University. Together, these collaborators have been investigating ways that cloud computing technologies can benefit the Unidata community, focusing on the use of the National Science Foundation's Jetstream cloud computing environment.

As part of this ongoing investigation, Julien and Unidata colleagues have been porting traditional Unidata technology offerings such as the Local Data Manager (LDM) and THREDDS Data Server (TDS) to the Jetstream environment. More recently, Julien has been working together with XSEDE ECSS colleagues to create a JupyterHub server that can accommodate multiple students in a workshop or classroom setting, enabling them to run and see results from their own interactive atmospheric science notebooks easily and with minimal setup or configuration.



Unidata Science Gateway on Jetstream



Building upon our previous containerization efforts, we are continuing to enhance the Unidata Science Gateway on NSF-funded XSEDE Jetstream Cloud: http://science-gateway.unidata.ucar.edu/. A collection of Unidata related technologies can be found here for our community to make use of directly or with client applications such as the IDV. The following resources are available on this gateway:

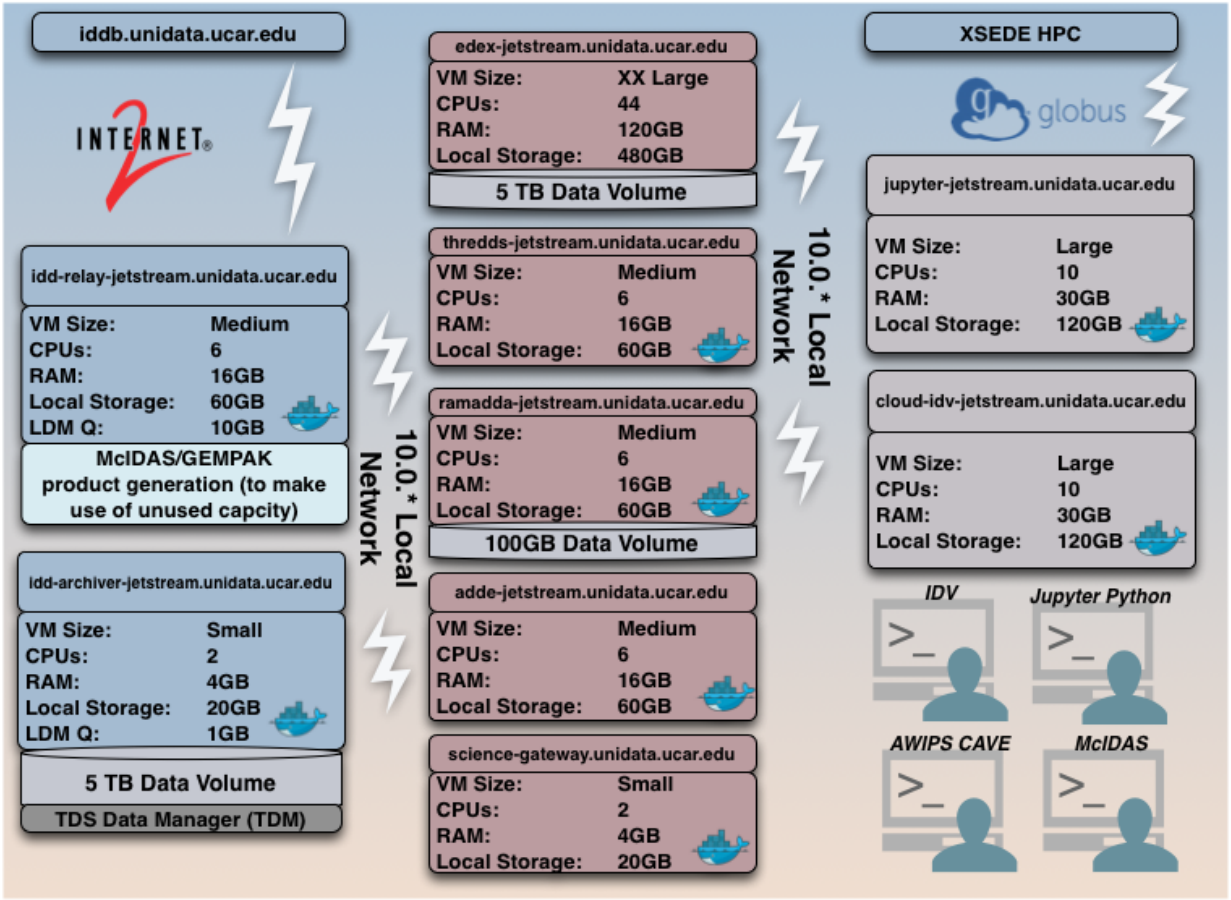
- An experimental JupyterHub server containing Unidata Jupyter notebook projects.
- A TDS supplying a good portion of the data available on the IDD with a five day archive.
- EDEX server
- RAMADDA containing IDV bundles that retrieve data from Jetstream data servers.
- Two LDM nodes
- ADDE Server

- IDV Jetstream plugin that allows easy access to Jetstream installations of the TDS, RAMADDA and ADDE from the IDV.

Gateway users, coupled with XSEDE HPC resources, can achieve complete end-to-end scientific computing workflows. In the past six months, we have done five presentations on this work:

- EGU General Assembly 2018 | April 8–13, 2018 - Vienna, Austria
- OGC TC/PC Meetings | June 4-8, 2018 - Fort Collins, Colorado, USA
- IWSG 2018 | June 13-15 2018 - Edinburgh, Scotland
- PEARC 18 | July 22-27, 2018 - Pittsburgh, PA USA
- ESIP Summer Meeting 2018 | July 17-20, 2018 – Tucson, AZ USA

A complete bibliography of this effort is available here.



Dependencies, challenges, problems, and risks include:

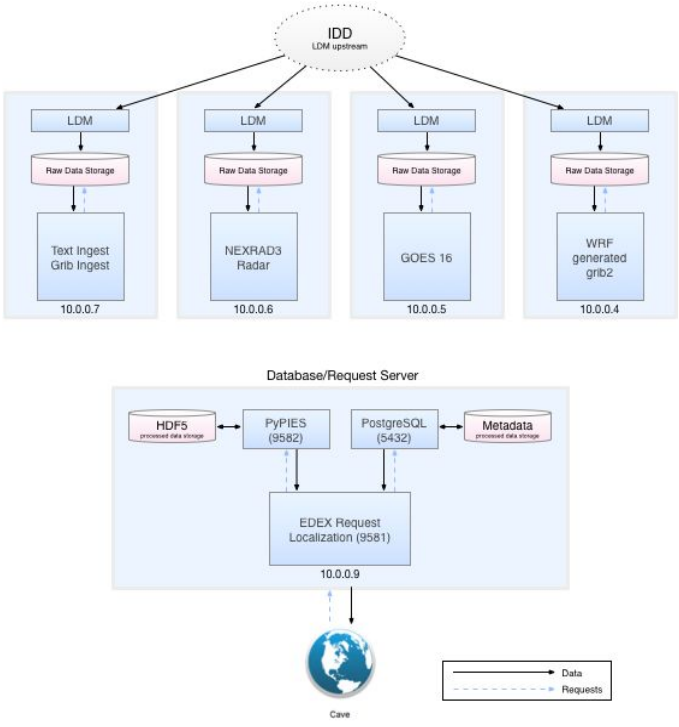
- We would like to transition from an experimental and research and development approach, to an operational mode. To achieve this objective, we are working with the Jetstream team to improve VM uptime availability. To that end, Unidata has

ameliorated its monitoring of Jetstream VMs so that we have accurate metrics concerning VM availability that we can communicate to the Jetstream team.
- Unidata technical staff continues to work closely with Unidata system administrators to ensure cloud VMs (especially on Jetstream) are adhering to Unidata security standards.

AWIPS EDEX in the Cloud

Unidata continues to maintain an EDEX data server on the Jetstream cloud, serving real-time AWIPS data to CAVE clients and through the python-awips data access framework (API). The distributed architectural concepts of AWIPS allow us to easily scale EDEX in the cloud to account for the size of incoming data feeds. By isolating the database/request processes to a single machine, we avoid *data serving* competing with *data decoding* on the same machine, minimizing the chance of reaching system memory limits which can result in EDEX shutdown.

We continue work using Jetstream to develop cloud-deployable AWIPS instances, both as imaged virtual machines (VMIs) available to users of *Atmosphere* and *OpenStack*, and as docker containers available on Docker Hub and deployable (soon) with the *xsede-jetstream* toolset.



Nexus Server on Jetstream

Unidata continues to run a Nexus Server on Jetstream for the distribution of netCDF-Java artifacts (e.g., `netcdfAll.jar`, `toolsUI.jar`, `ncIdv.jar`): https://artifacts.unidata.ucar.edu. netCDF-Java documentation is also hosted at that location.

Docker Containerization of Unidata Technology

We continue to employ Docker container technology to streamline building, deploying, and running Unidata technology offerings in cloud-based environments. Specifically, we are refining and improving Docker images for the IDV, LDM, ADDE, RAMADDA, THREDDS, AWIPS, and Python with Unidata Technologies. In addition, we also maintain a security-hardened Unidata Tomcat container inherited by the RAMADDA and THREDDS containers. Independently, this Tomcat container has gained use in the geoscience community.

Progress has been made on the following:

These containers are in active use and are maintained where needed. For example, we recently updated the base Tomcat images to incorporate new versions of Tomcat.

Dependencies, challenges, problems, and risks include:

It is unlikely that most of our community will 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. For example, on Jetstream we have a JupyterHub server currently in development: https://jupyter-jetstream.unidata.ucar.edu. This server was deployed with the aid of cloud computing technologies including Docker. These details, however, are hidden from the user.

In addition, there are overlapping (perhaps, competing or complementary) technologies such as Ansible that are emerging alongside Docker that need to be investigated.

Ongoing Activities

Amazon Web Service Activities and NOAA Big Data Project

NOAA Big Data Project

- In collaboration with Unidata, NOAA is delivering 20+ years of NEXRAD Level II data via Amazon Web Services. LDM and THREDDS Data Server (TDS) software are being employed to deliver these data.
- Started transferring GOES-16 data to the Amazon cloud S3 bucket.
- TDS on AWS for level II NEXRAD (For .edu access only): http://thredds-aws.unidata.ucar.edu/thredds/catalog.html
- AWS Explorer (Public access): https://s3.amazonaws.com/noaa-nexrad-level2/index.html
- Public Bucket for level II NEXRAD: https://noaa-nexrad-level2.s3.amazonaws.com
- Continue to populate the NEXRAD level II archive with real time data.

- Continue to populate new GFS .25 degree output and NCEP HRRR output to an S3 bucket for access. We did not place a TDS on this collection as this output is available from our standard sources.
- Unidata continues to get requests from other UCAR/NCAR groups, to partner and lend assistance in cloud computing, especially in the AWS cloud.

Product Generation for IDD

For the past three years, Unidata generated products for the IDD, FNEXRAD and UNIWISC data streams have been created by a VM hosted in the Amazon cloud. This production generation has been proceeding very smoothly with almost no intervention from Unidata staff.

CloudIDV, CloudStream, Cloud Control

- We have released the technology enabling CloudIDV in a form that can be easily leveraged by other projects looking to bring legacy software to the Cloud. We are currently trying to build the CloudStream community via conference presentations and outreach.
- In addition, we continue to experiment with CloudIDV on the Jetstream Cloud. We are investigating CloudIDV for data-proximate visualization of the WRF-hydro modeling system.
- We presented at both AGU 2017 and AMS 2018 on CloudIDV and CloudStream.

Open Commons Consortium Award

The Open Science Data Cloud, a resource of the Open Commons Consortium (OCC), provides the scientific community with resources for storing, sharing, and analyzing terabyte and petabyte-scale scientific datasets. The OSDC is a data science ecosystem in which researchers can house and share their own scientific data, access complimentary public datasets, build and share customized virtual machines with whatever tools necessary to analyze their data, and perform the analysis to answer their research questions. Unidata is a beta user of resources in the Open Science Data Cloud ecosystem and we have been provided cloud-computing resources on the Griffin cloud platform. Our allocations are renewed on a quarterly basis and Unidata is partnering with OCC on the NOAA Big Data Project. Given the limited staff resources and many ongoing cloud activities on AWS, Azure, and XSEDE environments, Unidata's activities on the OSDC have been in a temporary hiatus. We are hoping to ramp up our OSDC efforts in the upcoming months.

New Activities

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

Unidata Science Gateway

Forthcoming Presentations

- AMS Annual Meeting 2019 | January 6–10, 2019 - Phoenix, AZ

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

Unidata Science Gateway

We would like to promote and advertise the science gateway (http://science-gateway.unidata.ucar.edu/) to our community.

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

Unidata Transitioning to the Cloud

In the long-term, we would like to explore the possibility of migrating some core Unidata services onto the cloud.

Relevant Metrics

Docker image downloads are available from Unidata's Dockerhub repository.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Making Unidata data streams available via various commercial (e.g., Amazon) and not-for-profit (e.g., NSF XSEDE Jetstream) cloud services will allow our community to access data quickly and at low or even no cost. Moreover, our users can benefit from high data bandwidth capability provided by various cloud computing platforms, and in some cases, Internet2 capability. Lastly, cloud computing offers the possibility of accessing geoscience data in a "data-proximate" manner where users can perform analysis and visualization on, at times, unwieldy data sets next to where the data reside.

2. **Develop and provide open-source tools for effective use of geoscience data**
   Containerization technology complements and enhances Unidata technology offerings in an open source manner. Unidata experts install, configure and in some cases, security harden Unidata software in containers defined by Dockerfiles. In turn, these containers can be easily deployed on cloud computing VMs by Unidata staff or community members that may have access to cloud-computing resources. Unidata staff develop Docker containers in an open-source manner by employing software carpentry best-practices and distributed version control technology such as git.

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

4. **Build, support, and advocate for the diverse geoscience community**
   Transitioning Unidata technology to a cloud computing environment will increase data availability to new audiences thereby creating new and diverse geoscience communities.

Prepared *September 2018*

Status Report: Community Services

April 2018 - September 2018

Doug Dirks, Jeff Weber, Joshua Young

Areas for Committee Feedback

We are requesting your feedback on the following topics:

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:

- Summer 2018 Unidata Interns Wrap Up Their Projects
- NetCDF Licensing Remains Unchanged: Libraries are Open-Source and Freely-Available
- Python-Focused Software Training Workshop at Jackson State University
- Register Now for the 2018 Software Training Workshops
- 2018 Users Workshop Explores Evolving Geoscience Workflows
- What's New in IDV 5.5
- 2018 Community Equipment Awards
- 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:

- Drafting the next core funding proposal to the NSF.
- Planning for participation in upcoming professional society conferences (2018 AGU Fall Meeting and 2019 AMS Annual Meeting).
- Engagement with CUAHSI to support the NFIE and WRF-Hydro at the NWC.

- Assisting the EarthCube Science Support Office staff with back-office tasks.
- Continue to serve on the CUAHSI HIS standing committee.
- We continue to update Unidata's social media channels (Facebook, Twitter, Google+).
- We continue to publish short videos/screencasts on the Unidata YouTube channel.
- We continue to actively support the NCAR/SOARS program.
- Actively participate in Super Science Saturday.
- Engage and support the Undergraduate Leadership Workshop (ULW) at UCAR.
- Support the development and operation of the UCAR:NCAR Equity and InclusiON (UNEION) community of practice.

Dependencies, challenges, problems, and risks include:

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

Ongoing Activities

We plan to continue the following activities:

- Event planning activities for the Users Workshop in June will continue (and increase)
- Work on the Unidata core funding proposal to NSF
- Engagement with EarthCube, ESIP, and science or cyber communities at large
- NAWIPS migration to AWIPS, including the overall AWIPS project
- Ongoing development of news articles for publication through News@Unidata
- Continue to support and contribute to governing committees
- Seminars
- Outreach
- Inclusion and equity
- Engagement with professional societies
- Support for cloud-related projects
- Further development of the Data Management Resource Center
- Support the pursuit of funding
- Site visits as the budget allows
- Engage other UCAR/NCAR divisions regarding Unidata software use i.e. CESM/IDV

New Activities

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

- Expanded emphasis on cloud-related activities
- Support the Users Committee's work towards planning the next Triennial

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

- Continue to engage the hydrologic community regarding WRF-Hydro/IDV interactions and the National Water Center's efforts
- Seek additional opportunities to engage and listen to the community

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

- Provide additional data management and cloud-related training

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:

- 53,919 unique pageviews (56,470 in previous period)
- 10.1% of total unique pageviews 10.1% in previous period)

Top community pages

1. All blog pages
   43063 unique pageviews (42422 in previous period)
   71% of total community pageviews (79% in previous period)
2. www.unidata.ucar.edu/events
   11625 unique pageviews (3888 in previous period)
   19% of total community pageviews (7% in previous period)
3. www.unidata.ucar.edu/community
   2800 unique pageviews (4678 in previous period)
   5% of total community pageviews (9% in previous period)
4. www.unidata.ucar.edu/about
   2397 unique pageviews (2393 in previous period)
   4% of total community pageviews (4% in previous period)

Social media statistics, March 15, 2018

1. # of Twitter followers: 904 (up from 824)
2. # of Facebook followers: 754 (up from 717)

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 *August 2018*

Status Report: Community Equipment Awards

Sponsored by the National Science Foundation

April 2018 - September 2018

Admin Group

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Possible theme(s) for the 2019 Unidata Community Equipment Awards;
2. Please consider volunteering to serve on the 2019 Review Panel;
3. Suggestions from previous panel members on how to improve the program

Community Equipment Awards

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, in conjunction with the Triennial Users Workshop, special consideration was given to proposals that included one or more of the following:

- Projects that advance data-proximate analysis of large remote datasets (e.g. GOES-16 satellite data).
- Projects that facilitate the use of machine learning techniques and data analytics.
- Installation of equipment that provides student access to and use of GOES-R series satellite data.
- Installation of equipment for the operational use of ensemble models and the development of those models' predictions to share with the broader geoscience community.

A Request for Proposals was sent out on December 19, 2017 with a March 16, 2018 submission deadline. The Review Panel met on April 5 at the Unidata Program Center and recommended that the following proposals be funded:

- Jackson State University, Dr. Duanjun Lu, "Upgrade the JSU Meteorology Computing Lab by Installation of AWIPS II EDEX Server and CAVE Clients" proposal
- Pennsylvania State University, Dr. George Young, "Upgrading the Penn State IDD Relay for the Next Generation" proposal

- San Jose State University, Alison Bridger, “San Jose State University Unidata Equipment Proposal” proposal
- Texas Tech University, Eric Bruning, “Cloud-ready Processing and Dissemination of GOES-16 Geostationary Lightning Mapper Gridded Imagery” proposal
- University of Nebraska - Lincoln, Natalie Umphlett, “Enhanced Accessibility of Climate Data for Research and Teaching through a THREDDS Data Server” proposal
- University of Wisconsin-Madison, Gregory Tripoli, “A Community THREDDS/ADDE Data Server and IDD Infrastructure Upgrade at UW-Madison” proposal

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

Relevant Metrics

Since taking over the management and administration of the Equipment Awards program in 2003 on behalf of the NSF, Unidata has made 98 awards totaling over $1,200,000.

Prepared August 2018

Status Report: GOES-R Class Satellites

April 2018 - September 2018

Mike Schmidt, Tom Yoksas

Questions for Committee Members

- What coverages and increased spatial and temporal resolutions should be considered for UNIWISC IDD feed imagery?

  For instance, should coverages formerly provided by the NWS via NOAAPort (e.g., Super National, hemispheric, etc.) be added to the UNIWISC IDD feed?

- Which GOES-16 bands (wavelength channels) and coverages should be converted from netCDF-4 to McIDAS AREA format so that GEMPAK users will continue to be able to use currently satellite imagery?

  At its April, 2018 meeting, the User's Committee requested that Full Disk, CONUS and Mesoscale coverages for the high resolution visible (0.64 um), mid level water vapor (6.93 um) and thermal infrared (10.3 um) be made available in the UNIWISC feed. Are there other wavelength channels and coverages (sectors) that should also be made available?

- What kind(s) of data access methods are most desired/usable for the community?

- Other questions?

Activities Since the Last Status Report

- Began ingesting all GOES-17 products available in the GOES ReBroadcast (GRB) on Tuesday, August 28, the day after ABI imagery delivery was enabled in the G17 GRB

  Ingest is being done on the 4.5 m satellite dish that was installed at the Mesa Lab using funds provided by NOAA's GOES-R office. GOES-16 ingest was moved to a 3.8 m satellite dish located outside of the cafeteria located in Foothills Lab building 2.

- Setup a test IDD feed (SPARE) for the GOES-17 products the same day as ingest was begun.

  We are currently tuning this capability and will be making it available to sites that are eager to test GOES-17 products.

- Setup ADDE serving of the test GOES-17 ABI imagery on the same machine that ingesting the data

We will allow sites that want to test the imagery access to the ADDE server on a case-by-case basis.

- Preparing to implement UW/SSEC's "fanout server" (redistribution of the GRB-200 UDP unicast stream over TCP) for GOES-17 products

  We will be feeding from one of the SSEC GOES-17 fanout servers, and they are already feeding from the ingest machine that we operate.

- Continuing to deal with Terrestrial Interference (TI) on the satellite dish we installed at the NCAR Mesa Lab

  In the fall of 2017 we began experiencing significant TI in the GOES-16 signal being received by the satellite dish at the NCAR Mesa Lab. An outcome of the discussions we had with Quorum Communications (the manufacturer of the electronics we use in our GOES-R/S ingest installations) was our moving of the GOES-16 ingest to a 3.8 m satellite dish located at the UCAR FL-2 location. The relocation of GOES-16 ingest required that an additional signal cable be pulled from the satellite dish that was repurposed from GOES GVAR ingest into the 2nd floor FL-2 NCAR/RAL computer room where our ingest electronics are located. The cost of this work was contributed by the UCAR/NCAR networking group.

  GOES-16 ingest on the FL-2 satellite dish has been working well except during periods where maintenance is being done in the tree-lined plaza between FL-1 and FL-2. The worst interference is experienced when cleaning crews use gas-powered leaf blowers in the plaza. A secondary source of "interference" (signal degradation is a better description) is the trees that fill the plaza. We expect that signal levels will improve when the leaves fall off of the trees sometime later this month. In the long term, the trees in the plaza will be removed (by UCAR) to facilitate construction to resolve drainage issues. Afterwards, some trees may be restored to the area, but we expect they will be sized and placed to avoid future problems

Ongoing Activities

We plan to continue the following activities:

- Ingest the GOES ReBroadcast (GRB) from GOES-16 in real-time using a 3.8 m satellite dish located outside of the cafeteria at FL-2

- Ingest GOES-17 products using the 4.5 m satellite dish that we installed on the eastern pad at the NCAR Mesa Lab using the NOAA funded UW/SSEC/CIMSS Community Satellite Processing Package (CSPP) for Geostationary Data (GEO) package

  We anticipate that TI we have been experiencing on the Mesa Lab dish will continue to decrease as the dish is swung to point at the GRB GOES-West location of 137 W longitude. The reason for this belief is the great reduction in TI that we observed

when repointing the dish from the GOES-East location of 75 W to the GOES-17 test location at 89 W.

- Continue to distribute GOES-16 data via the LDM/IDD and serve the data via McIDAS ADDE, the TDS and AWIPS EDEX

  We will also be making GOES-17 data available in the same ways when it becomes GOES-West in the late fall/early winter this year or possibly sooner if the demand for early access to the data is great enough.

- Continue working with NCAR/EOL to support their CSPP GEO installation at the NCAR Mesa Lab

    - This installation provides an additional institutional redundancy for GOES-16 ingest

- Canvas the community to learn more about their GOES-16/17 data needs

- Investigate additions to the IDD CONDUIT data stream that would be useful for creation of new GOES-16/17 based Level 2 products

Future Activities

- Continue working with SSEC on their *fanout* approach that insulates GRB ingestion from expected (e.g., NCAR twice per year power downs; twice per year solar interference periods; etc.) and unexpected (e.g., TI caused) service interruptions

- Unidata-Wisconsin (UNIWISC) IDD imagery was updated to replace GOES-13 imagery with GOES-16 imagery in GOES-East sectors in early January, 2018.

  Since the middle of May, we have been converting Full Disk, CONUS and Mesoscale coverages for three wavelength channels (0.64 um VIS, 6.93 um WV and 10.3 um IR) into McIDAS AREA format as a test for changes to the Unidata-Wisconsin (IDD feed type UNIWISC) to be rolled out this month.

- Distribute value added GLM products (netCDF4 format) created by Eric Bruning of Texas Tech via the LDM/IDD and serve the same data via McIDAS ADDE, the TDS and AWIPS EDEX

- Establish a test bed for the creation of Level 2 (L2) products from GOES-16/17 imagery, model output and observational data

  The objective is to provide the capability of running user site submitted algorithms to create L2 products and make them available for testing for a short period of time via the IDD, McIDAS ADDE, the TDS and AWIPS EDEX.

Relevant Metrics

- Lots O Data!

  The volume of GOES-16 imagery, space weather and GLM products can be seen in the real-time statistics plot from our GOES-R ingest machine:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?SATELLITE+ingest.unidata.ucar.edu

  The volume of GOES-17 imagery, space weather and GLM products can be seen in the real-time statistics plot from our GOES-S ingest machine:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?SPARE+cluster5.unidata.ucar.edu

  For comparison, the volume of just the GOES-16 ABI imagery and GLM Level 2 products can be seen in the real-time statistics plot from NCAR/RAL's ingest machine:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?EXP+satops3.rap.ucar.edu

- Feeding data to a growing list of sites via the IDD:

  We are distributing all or part of the GOES-16 GRB products to:

  - Groups within UCAR/NCAR (2: all products (Unidata) or ABI imagery and GLM L2 products (RAL))
  - U.S. Universities (21: variety of feeds)
  - U.S. Government (2: all products to 2 NOAA sites)
  - International (2: Full Disk imagery and GLM L2 products)

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Standing up ADDE and TDS data services for real-time GOES-R/S data will benefit the greater Unidata community.

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 bundled in AWIPS.

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.

Prepared *September 2018*

Status Report: Internet Data Distribution

April 2018 - September 2018

Steve Emmerson, Mike Schmidt, Jeff Weber, Tom Yoksas

Questions for Committee Members

- Suggestions regarding content of data streams like CONDUIT, FNEXRAD, UNIWISC and LIGHTNING? ... We (UPC, the Unidata community and UAlbany for the NLDN component of LIGHTNING) have control of the content of these data streams, so they are open for suggestions.
- The top level IDD relay clusters we maintain were upgraded to support relay of the full volume of GOES-16/7 imagery and products being received via redundant GOES-16/7 downlinks. The data volume for this feed can be seen in the **SATELLITE** (aka **DIFAX**) feed in the volume listing for a variety of machines operated by Unidata. The question for the governing committees continue to be if real-time delivery of GOES-16/7 data via the LDM/IDD is still desired, **and** if end typical users can handle the data volume.

Activities Since the Last Status Report

Internet Data Distribution (IDD)

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

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

Maximum hourly volume  91623.724 M bytes/hour
Average hourly volume  61372.022 M bytes/hour

Average products per hour     485179 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)             (M byte/hour)   number/hour
FSL2                    10766.490    [ 17.543%]    19966.651    13229.021
NEXRAD2                  9321.264    [ 15.188%]    12533.394    90753.957
NGRID                    8514.015    [ 13.873%]    13200.785    62352.915
CONDUIT                  7915.147    [ 12.897%]    20745.264    90269.851
DIFAX                    6426.086    [ 10.471%]     8504.604     3288.426
NOTHER                   5920.507    [  9.647%]     9385.784     9915.085
EXP                      5449.378    [  8.879%]     8854.779     5222.915
NEXRAD3                  2897.408    [  4.721%]     3704.448   116056.426
FNMOC                    2300.405    [  3.748%]     9255.044     3335.128
HDS                      1221.580    [  1.990%]     1597.422    42196.553
```

```
GEM                    165.516    [  0.270%]     1004.481      995.745
FNEXRAD                138.292    [  0.225%]      167.328      104.596
UNIWISC                 98.986    [  0.161%]      143.621       49.872
NIMAGE                  94.914    [  0.155%]      143.319      122.128
IDS|DDPLUS              76.791    [  0.125%]       91.231    46573.830
LIGHTNING               65.244    [  0.106%]      102.689      712.447
```

The 6-8 GB/hr value in IDD data volume shown as **DIFAX** represents the ingestion of GOES-16 data, and this will effectively double when GOES-17 data is added to the IDD.

New Data Distribution: Unidata is taking over the data distribution of GPS radio occultation solutions from COSMIC. COSMIC will still gather incoming GPS data and create the solutions, but due to hardware constraints COSMIC has requested Unidata to provide distribution from our top level IDD relay clusters (idd.unidata.ucar.edu and iddb.unidata.ucar.edu) to the community. The solutions (Precipitable Water Vapor and Total Electron Content-Ionosphere) are in netCDF format and are available in the **GPS** feedtype.

Ongoing Activities

We plan to continue the following activities:

- Unidata receives the NOAA/GSD experimental 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 of university sites on **hrrr.unidata.ucar.edu**.
- The NCEP operational HRRR is being served with other model output at:

  http://thredds-jumbo.unidata.ucar.edu/thredds/catalog/idd/forecastModels.html (.xml for machine access)

- Other data sets we continue to explore with NOAA/GSD/ESRL are:
  - FIM
  - HIWPP
- NCEP (operational) HRRR fields and forecasts times have been added to the CONDUIT IDD datastream.

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 UCAR NOAAPort dish pointing was moved from SES-1 (approx. 101 W) to Galaxy 23 (approx. 89.5 W) in December 2017. Since the move, the indicated Carrier to Noises/EsNos reported by Unidata Novra S300N receivers have improved from the mid-15s to mid-17s with peaks occasionally being higher than 18. EsNos in this range indicates exceptionally good reception quality.

- The NOAAPort-derived data streams (**HDS**, **IDS|DDPLUS**, **NGRID**, **NIMAGE**, **NEXRAD3** and **NOTHER**) are being redundantly injected into the IDD at three geographically separate locations: Unidata, UW/SSEC, and LSU/SRCC. The **NOTHER** data stream contains GOES-16 and GOES-17 tiles that need to be stitched together to make full image scenes usable to end-user applications. Unidata provides Ryan May’s ldm-alchemy package in the Unidata section of Github for this purpose.

  We continue to look for a fourth ingest site to increase robustness of the IDD distribution of NOAAPort derived data.

- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.13.6, but will soon be updated to v6.13.7.

- We recently discovered that the LDM/IDD Product IDs for the GOES-16 Level 2 (L2) products that are being distributed in NOAAPort are not sufficiently descriptive to easily allow for saving of all products to disk via LDM pattern-action file actions.

  We are testing use of the product sequence number (assigned by our NOAAPort ingest code) in pattern-action file actions to work around this limitation.

Relevant Metrics

- Approximately **573** machines at **243** sites are running LDM-6 **and** reporting real-time statistics to the UPC.

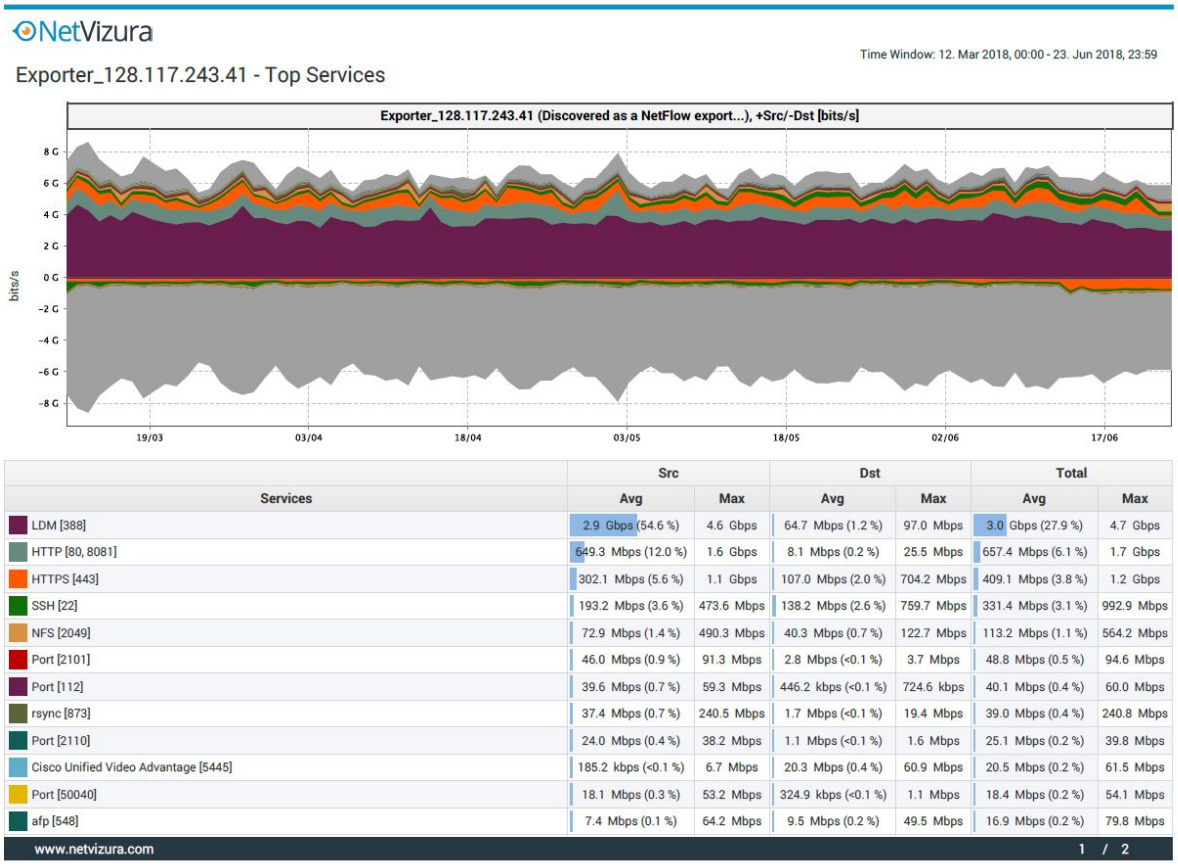
  Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

- 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 relays, **idd.unidata.ucar.edu and iddb.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.

Over the period from March 12, 2018 through June 23, 2018 the average volume of LDM/IDD data flowing from the UCAR/NCAR network averaged around 3.0 Gbps (~32.4 TB/day), and peak rates reached 7.0 Gbps (which would be ~75.6 TB/day if the rate was sustained).



Services	Src Avg	Src Max	Dst Avg	Dst Max	Total Avg	Total Max
LDM [388]	2.9 Gbps (54.6 %)	4.6 Gbps	64.7 Mbps (1.2 %)	97.0 Mbps	3.0 Gbps (27.9 %)	4.7 Gbps
HTTP [80, 8081]	649.3 Mbps (12.0 %)	1.6 Gbps	8.1 Mbps (0.2 %)	25.5 Mbps	657.4 Mbps (6.1 %)	1.7 Gbps
HTTPS [443]	302.1 Mbps (5.6 %)	1.1 Gbps	107.0 Mbps (2.0 %)	704.2 Mbps	409.1 Mbps (3.8 %)	1.2 Gbps
SSH [22]	193.2 Mbps (3.6 %)	473.6 Mbps	138.2 Mbps (2.6 %)	759.7 Mbps	331.4 Mbps (3.1 %)	992.9 Mbps
NFS [2049]	72.9 Mbps (1.4 %)	490.3 Mbps	40.3 Mbps (0.7 %)	122.7 Mbps	113.2 Mbps (1.1 %)	564.2 Mbps
Port [2101]	46.0 Mbps (0.9 %)	91.3 Mbps	2.8 Mbps (<0.1 %)	3.7 Mbps	48.8 Mbps (0.5 %)	94.6 Mbps
Port [112]	39.6 Mbps (0.7 %)	59.3 Mbps	446.2 kbps (<0.1 %)	724.6 kbps	40.1 Mbps (0.4 %)	60.0 Mbps
rsync [873]	37.4 Mbps (0.7 %)	240.5 Mbps	1.7 Mbps (<0.1 %)	19.4 Mbps	39.0 Mbps (0.4 %)	240.8 Mbps
Port [2110]	24.0 Mbps (0.4 %)	38.2 Mbps	1.1 Mbps (<0.1 %)	1.6 Mbps	25.1 Mbps (0.2 %)	39.8 Mbps
Cisco Unified Video Advantage [5445]	185.2 kbps (<0.1 %)	6.7 Mbps	20.3 Mbps (0.4 %)	60.9 Mbps	20.5 Mbps (0.2 %)	61.5 Mbps
Port [50040]	18.1 Mbps (0.3 %)	53.2 Mbps	324.9 kbps (<0.1 %)	1.1 Mbps	18.4 Mbps (0.2 %)	54.1 Mbps
afp [548]	7.4 Mbps (0.1 %)	64.2 Mbps	9.5 Mbps (0.2 %)	49.5 Mbps	16.9 Mbps (0.2 %)	79.8 Mbps

www.netvizura.com 1 / 2

Cluster real server backends and accumulator nodes routinely have instantaneous output volumes that approach 2 Gbps. Bonding of pairs of gigabit Ethernet interfaces was needed to be able to support these output data rates. The next generation of cluster machines to be installed soon will have 10 Gbps Ethernet capability.

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 bundled in AWIPS.

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 actively pursuing IDD relay in the Latin America based IDD-Caribe.

Prepared *September 2018*

Status Report: IDV with RAMADDA

April 2018 - September 2018

Yuan Ho, Cece Hedrick, Julien Chastang

Areas for Committee Feedback

We have no questions at this time.

Activities Since the Last Status Report

IDV System Changes

__Latest netCDF-Java Version__

The version of the netCDF-Java library currently distributed with the latest stable IDV version (5.5) is the 4.6.11-SNAPSHOT . The prior version of netCDF-Java to be distributed with the IDV was 4.6.10. There have been many improvements and bug fixes in that range. The complete release notes for these versions can be found here.

__ISL Changes__

ISL movie and image capture improvements for DRILSDOWN project.

__Java Version__

The IDV is distributed with Java 8 u51. Because of a Windows 10 related bug on some older Intel hardware, this will be the latest version of Java we will be deploying with the IDV.

__Java 3D Version__

The IDV is distributed with Java 3D version 1.6.0 Final. Java 3D 1.7 prerelease is under development and continues to be supported in some capacity by the open source community.

__HTTP to HTTPS__

As the web transitions from HTTP to HTTPS, we update URL references in the IDV (rbi files, plugin files, JNLPS, etc.) to reflect these changes.

__Java Code Refactorings__

Improving Java coding standards in the IDV code base.

__IDV Certifcates__

Java Webstart and Windows app certificates have been renewed and will be valid until May 30, 2020.

IDV Display Changes

__3D Isentropic Analysis__

Isentropic analysis is a very common and effective way to diagnose and visualize vertical motion of the atmosphere, the new isentropic analysis on the constant potential temperature surface helps us better understand the dynamic and stability of weather phenomenons. This new feature allows IDV users to do the isentropic analysis on a regular isobaric gridded dataset without any vertical coordinate conversion.

__GOES 16 GLM Display__

The GLM L2 product consists of geo-located and time-stamped events, groups, and flashes, with associated calibrated optical amplitudes (in units of Joules). The IDV has three new data source types: GLM EVENT Data files, GLM GROUP Data files, and GLM FLASH Data file to view these three GLM products.

__Scatter Analysis__

The IDV constructs scatterplots to illustrate different types of direction and strength of relationships of different weather parameters, where the values of one variable are along X axis and the values of the second variable are along the Y axis. The new scatter analysis is used to create a scatter plot to study the relationships between two 2D parameters. The Scatter analysis in the IDV also provides statistical information about the data as a whole, as well as the subset regions of the data.

__Grid 3D Surface Smoothing__

IDV extends the all 2D horizontal smoothing types to 3D gridded parameters. The smoothing is now available on 3D displays like isosurfaces of gridded dataset and level2 radar dataset.

__Latest Version of VisAD__

The SSEC team at UW, Madison has made a number of improvements to support 3D trajectories.

IDV WRF-Hydro Collaboration

Jeff Weber and Yuan are collaborating with David Gochis, and David Yates from NCAR-RAL assisting them ensuring WRF-Hydro data is CF compliant regarding Point Feature type. In

addition, we are helping them with visualization of this dataset in the IDV. Yuan will present to the National Water Center, on March 27th, findings regarding data format and ways to make the WRF-Hydro output more accessible by the Unidata community.

IDV Release

IDV 5.5 was released in May of 2018.

IDV EarthCube Proposal Awarded

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

Unidata hired Celia Hedrick, a Software Engineer II to work on this project.

IDV Publication Highlights

Synoptic–Dynamic Meteorology in 3D: Introducing an IDV-Based Lab Manual by Gary Lackmann, B. Mapes and K. Tyle

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

IDV and RAMADDA Training, Conference Attendance and Presence

__2018 American Meteorological Conference (AMS) Annual Meeting__

- The New Trajectory Display in the UNIDATA's IDV
- Synoptic–Dynamic Meteorology in 3D: Introducing an IDV-Based Lab Manual
- Drilling down from Python Statistical Analyses to Rich Interactive Case Study Visualizations, within Jupyter Notebooks

Ongoing Activities

We plan to continue the following activities:

__IDV scatter display__

We will enhance the scatter display feature of the IDV and cover 3D gridded dataset at each level, point feature dataset, and radial type dataset.

__GOES_16 GLM ADDE__

We will develop new adapter in the IDV to communicate with the ADDE server to access the GLM data available in the ADDE server.

__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.

New Activities

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

Relevant Metrics

__E-Support__

The IDV team continues to provide the geoscience community with high-quality support through e-support software and idv-users mail list. In the last half year the IDV team has closed ~30 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 113 “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 16,265 times compared with 15,000 from six months ago.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   RAMADDA allows geoscience specialists the ability to search and publish their IDV bundles on-line. Unidata's RAMADDA installation enables the IDV team to communicate more effectively to our users concerning their IDV issues. Specifically, during support ticket conversations, the IDV team requests that users upload pertinent data to RAMADDA for analysis. One of RAMADDA’s best features is the ability to upload a CDM file and obtain the OpenDAP link from the new entry. The DAP link can be shared and opened in the IDV. RAMADDA also has “server-side view” capability where a specific part of the file system can be made available through the RAMADDA interface. This feature can be helpful to view LDM data feeds, for example. The IDV team also takes advantage of RAMADDA to share instructional IDV screencasts with users.

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

Status Report: Information Technology

April 2018 - September 2018

Mike Schmidt, Matt Perna

Major Activities

IDD cluster -- We've completed the purchase of updated IDD cluster hardware for eventual installation at the NCAR Wyoming SuperComputing Center (NWSC). After configuration and installation at the NWSC, the cluster will replace the aging Foothills Lab (FL) cluster that has served us well for many years. Notable changes include 10Gb/s network service and the ability for LDM queue sizes to triple from the current 100GB.

GOES-16 -- We've moved the GOES-16 ingest dish from the Mesa Lab (ML) to the Foothills Lab (FL) cafeteria dish.. This resolved the terrestrial interference (TI) issue encountered with the ML dish. On the down side, we endure a significant signal degradation to complete signal loss when Facilities uses leaf blowers in the courtyard.

GOES-17 -- The ML dish is now actively receiving GOES-17 data and we have a system processing the data via CSPP-GEO on a system in the ML data center.

100Gb/s networking -- as UCAR upgrades their backbone infrastructure from 10-20Gb/s links to 100Gb/s links, we will continue to upgrade our data movers (IDD cluster nodes, data aggregators, motherlode clones) from 2 x 1Gbp/s bonded ethernet to 10Gb/s as necessary.. The addition of GOES-S data in late 2018 will require the need soon anyway.

UCAR FISMA -- UCAR currently has contractual FISMA requirements for a small parts of the organization (none for Unidata currently), but the plan is to start down the path of making the entire organization FISMA compliant at some level yet to be determined (probably low, possibly medium). We continue to attend meetings to stay connected to the process

Security -- we continue efforts to keep services and systems secure which takes consistent attention and occasional herculean efforts (to patch everything all at once).

LDM 7 node -- we maintain a LDM7 test node at the Front Range GigaPOP (FRGP) just off downtown Denver in co-location with the major backbone networks supporting FRPG participants (UCAR, ..). We expect to support intensive data movement and LDM testing for the next few years on this effort.

Desktop backups -- we're in the middle of evaluating an improved desktop backup solution

Ongoing Activities

We plan to continue the following activities:

- Day-to-day system and network support to the community as needed
- Resolve daily staff help desk issues
- Maintain security profile and exceed UCAR security standards

---

Prepared *September 2018*

Status Report: LDM

April 2018 - September 2018

Steve Emmerson, Tom Yoksas, Mike Schmidt, Yuanlong Tan (UVA)

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:

- Added missing parameters and vertical coordinates to GEMPAK's GRIB2 tables
- Improved handling of missing GEMPAK GRIB2 entries
- Improved performance of GEMPAK parameter search from O(n) to O(log(n))
- Improved logging
- Improved handling of filename patterns by plotMetrics(1)
- Added ability for plotMetrics(1) to create .png files
- Improved concurrency of "ldmadmin addmetrics"
- Created delayQueue(1) for use in GOES-16 ingest
- Completed unit-testing of integrated LDM7

Dependencies, challenges, problems, and risks include:

- Dealing with missing GEMPAK table entries represents a continuing problem area (not new)
- The LDM is sometimes held responsible for decisions made by the NWS on how to categorize data products (also not new)

Multicast LDM (aka LDM-7)

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

Another multi-year grant has been awarded by the NSF for this project. The funds were finally available at the end of September.

Progress has been made on the following:

- Integration of the Python code to create a multi-point VLAN on Internet2

- Integration of the Python code to create virtual circuits on Internet2
- Provisioning of LDM7 servers at University of Maryland, Rutgers, University of Utah, University of Washington, and University of California at San Diego

Ongoing Activities

We plan to continue the following activities:

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

New Activities

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

- Install LDM-7 servers at 5 more participating universities
- Collect metrics on the LDM7 network

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

- Continued development and deployment of LDM-7

Relevant Metrics

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

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

Status Report: McIDAS

April 2018 - September 2018

Tom Yoksas

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Are there any features that users would like to be added to Unidata McIDAS-X and/or ldm-mcidas?

Activities Since the Last Status Report

Aside from routine updates/bugfixes to existing code and tables, the main area of investigation has to been the creation of ADDE servers for NOAAPort-delivered GOES-16 satellite imagery.

Current Activities

- Unidata McIDAS version 2018 will be released this month

  v2018 includes all SSEC versions up to and including the current McIDAS-X and -XCD releases, both of which are v2018.1.

- McIDAS-X is used to convert GOES-16 ABI imagery that is in netCDF4 format to McIDAS AREA format that is usable by all supported display and analysis packages except Python/MetPy. Support for McIDAS AREA format is on the list of things to do for Python/MetPy.

The latest releases feature the following:

- Updated ADDE servers for Himawari and GOES-R imagery

  A notable feature of core McIDAS support for imagery from platforms such as GOES-R and Himawari is the implementation by SSEC of an Expanded Stretch Feature which includes bit-depth changes to increase the detail shown in certain visible, water vapor and short-wave infrared imagery.

- ADDE server for NOAAPort GOES-16/17 Sectorized Cloud and Moisture (SCMI) imagery

  The core McIDAS ADDE servers do not support the tiled GOES-16/17 image sectors that are being distributed in the NOAAPort SBN. ADDE servers have been developed in Unidata to support these NOAAPort-delivered images.

- Updated applications designed for GLM L2 lightning data

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 occasionally receives 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

- Bandwidth usage by the Unidata McIDAS ADDE servers routinely exceeds 28 TB/month for the period spanning January 1, 2018 through May 31, 2018..

  The amount of data served by Unidata ADDE instances increased substantially with the availability of GOES-16 ABI imagery and GLM lightning data. ADDE serving of GOES-16 GRB data has been steady at about 5 TB/week.

- McIDAS-X/-XCD Inquiry Metrics

ldm-mcidas Decoders Activities

Development

ldm-mcidas releases are made when needed to support changes in software development and operating system environments. **ldm-mcidas** v2012 was released at the end of September, 2012. This package is in the process of being updated to support changes to various data streams.

Geostationary Satellite Data Ingest and Data Serving

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-16) data via McIDAS ADDE has been averaging approximately 19 TB/month since GOES-16 assumed the duties of GOES-East.

- Direct, programmatic access to real-time GOES-West (GOES-15) data via McIDAS ADDE routinely exceeds 2 TB/month.

Planned Activities

Ongoing Activities

Continued ingest, distribution via the IDD and ADDE serving of GOES-West and GOES-14 imagery from the existing constellation of GOES GVAR platforms.

Continued ingest, distribution via the IDD and ADDE serving of GOES-East imagery from the new GOES GRB platform, GOES-16.

These efforts require maintenance of the satellite ingest and data serving equipment.

New Activities

Operationalize the ingest, distribution via the IDD and ADDE serving of GOES-17 imagery. IDD distribution of GOES-17 imagery has already been setup in a test mode, and is available to community members via a point-to-point IDD feed.

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 worldwide 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. 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 *September 2018*

Status Report: netCDF

April 2018 - September 2018

Ward Fisher, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. To what extent is Amazon S3 used within your organization? Would you benefit from native netCDF support for S3 storage?
2. Are there other cloud-based block storage formats/locations (zarr, Azure, etc) that are actively in use? That we should consider storing.
3. Are there any emergent avenues (stack overflow, etc) for user support which the netCDF team should investigate?
4. How can we encourage more user testing of the release candidates we provide?

Activities Since the Last Status Report

We are using 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 106 open issues for netCDF-C, 27 open issues for netCDF-Fortran, and 16 open issues for netCDF-C++. The netCDF Java interface is maintained by the Unidata CDM/TDS group 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:

- Addition of user-defined compression filters.
- Work towards enhanced parallel I/O.
- Further extension of the netCDF build-and-test platforms using Docker technology.
- Further enhancements to the netCDF documentation.
- Extended continuous integration platforms have been adopted.
- An architecture roadmap is available describing how the netcdf-c library will support thread-safe operation in *nix* and Windows environments. The draft proposal is available as netcdf-c github issue #382.
- We have seen an uptick in the number of contributions to the netCDF code base(s) from our community. While these contributions require careful review and consideration, it is encouraging to see this model of development (enabled by our move to GitHub) being more fully embraced by our community.
- Increased support for native CDF5 on 64-bit platforms, potentially obviating the need for the parallel netcdf library for reading/writing CDF5 files in serial fashion.

Dependencies, challenges, problems and risks include:

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

Ongoing Activities

We plan to continue the following activities:

- Provide support to a large worldwide community of netCDF developers and users.
- Continue development, maintenance, and testing of source code for multiple language libraries and generic netCDF utility programs.
- Improve organization of Doxygen-generated documentation for netCDF-C and Fortran libraries.
- Extend collaboration as opportunities arise, for increasing the efficiency of parallel netcdf-3 and netcdf-4.

New Activities

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

- Seek out, and prepare material for upcoming, conferences and other outreach opportunities.
- Work on reducing the defects reported by static analysis.
- Release the next versions of netCDF-C, netCDF-Fortran, netCDF-C++.
- Modernize the netCDF documentation to provide easy access to documentation for older versions of netCDF.
- Provide thread-safety for the netCDF C library.

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

- Release an official Windows port of the netCDF-Fortran and netCDF-C++ interfaces.
- 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.
- Implement support for Amazon S3 in the netCDF C library.

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

- Improve scalability to handle huge datasets and collections.
- Improve the efficiency of parallel netcdf3 and parallel netcdf4.

Relevant Metrics

There are currently about 189,774 lines of code (up from 183,700 lines of code) in the netCDF C library source. The Coverity estimate for defect density (the number of defects per thousand lines of code) in the netCDF C library source has been increased slightly from **0.7** six months ago to **0.71** 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**. The jump in defect density is a result of the addition of the **DAP4** code. As this is new code, the initial defects are still being worked out.

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:

- **897,000** for netCDF-3
- **836,000** for netCDF-4
- **510,000** for HDF5
- **331,00** for GRIB2

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

- **320** for netCDF-3
- **646** for netCDF-4
- **11,100** for HDF5
- **913** 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 *March 2018*

Status Report: Python

April 2018 - September 2018

Ryan May, John Leeman, Sean Arms, Julien Chastang, Michael James

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. How are our existing suite of training materials suiting your needs? Any noticeable holes?
2. What are the most useful functionalities in MetPy and Siphon for your needs? What do we do well?
3. Are there any additions you'd like to make to MetPy's or Siphon's roadmap?

Activities Since the Last Status Report

Python Training Efforts

Python training efforts continue to be a valuable portion of the Python portfolio. We continue to be successful in identifying opportunities to offer training within our resource constraints. Not only do these generate significant goodwill and grow our audience, but they are a significant source of information to inform our library development. One challenge is to balance time dedicated to workshop preparation and logistics against time devoted to support and Python software development.

Progress has been made on the following:

- John Leeman and Ryan May, together with our summer interns Jon Thielen and Hailey Johnson, taught a MetPy/Siphon workshop at Colorado State University in May to around 25 attendees
- John and Ryan taught a MetPy/Siphon workshop at Jackson State University 29-30 August 2018. The workshop was fairly well attended, with approximately 25 people, including several from the local National Weather Service forecast office.
- John and Ryan will be teaching a MetPy/Siphon workshop at San Jose State University on 4-5 October 2018.
- Ryan May and John Leeman, together with Kevin Goebbert, will be testing a short course using MetPy at the 2019 AMS Annual Meeting. The course will focus on practical application of MetPy (and Siphon) to synoptic meteorology.
- John Leeman continues to lead the "MetPy Mondays" effort. These weekly screencasts on the Unidata Developers' blog receive a lot of attention and feedback. Creating these also often uncover improvements for our software. MetPy Mondays have had two guest hosts, Kevin Goebbert and Jon Thielen, contributing videos on satellite data and XArray. We welcome additional community screencasts.

MetPy

MetPy continues to grow, both in features and in community. The volume of support requests continues to rise, with regular traffic across GitHub, E-Support/E-mail, and Stack Overflow. Development going forward will continue to be driven by requirements for our dedicated awards (in addition to bug reports and pull requests from community members). The primary efforts will be focused around improved units support, integration with xarray, data formats, and GEMPAK-like interface. The AMS 2019 talk on MetPy's GEMPAK-like interface has been accepted, and at that time we plan to debut the first iteration of this interface in MetPy 0.10.

As part of MetPy's NSF award, we have been tracking more in depth some engagement metrics. Highlights from these results include:

- MetPy was downloaded a total of 13300 times from Anaconda.org and the Python Package Index during the last year.
- 74 email support requests to support-python, which represents a large increase over previous years.
- During the last year the MetPy documentation web pages averaged 8100 page views per month, up from 4500 views per month the previous 6 months (the only time for which data are available). From these data it is clear that the most popular segments of the documentation are MetPy's examples (37% of all page hits) and the API documentation for calculations (15% of page hits).

We also conducted the first MetPy user survey during this period of performance, which indicated that overall MetPy is meeting the needs of its users. The results reflect the responses of 32 individuals. When asked "In your experience how easy is it to use MetPy for your various teaching and/or research activities?", 71.9% of respondents rated MetPy 4 or higher on a 5 point scale (with 1 being "not easy" and 5 being "very easy"); 25% of respondents gave a rating of 3, which indicates some room for improvement. When specifically asked about performing GEMPAK-like tasks, 84.4% of responses rated MetPy as "Good" or "Excellent", with 15.6% responding "Fair". When asked for specific enhancements, many responses mentioned cross-sections and additional grid diagnostics.

Progress has been made on the following:

- Community awareness continues to grow, with the volume of engagement and mentions on social media growing; the MetPy twitter account has reached 582 followers. Involvement and support requests continue to grow as well (e-support, Gitter chat), including quite a few contributions on GitHub.
- MetPy 0.9 (and 0.9.1) was released, including support for cross-sections, more GEMPAK calculations, and enhanced XArray integration.
- MetPy 0.8 was released. The headline feature of this release is dependence on XArray, automatic conversion of DataArrays for calculations, and support for parsing CF projection metadata to simplify plotting in CartoPy. There were also several new calculations as well as support for opening GINI files directly as XArray Datasets.
- Work towards requirements of MetPy-related NSF awards

Siphon and Data Processing

Siphon continues to grow and develop, though at a slower pace than MetPy; its development tends to be driven by obstacles to access of remote data. Siphon has begun to see an increase in community contributions, which is likely due in part to the inclusion of data sources beyond TDS.

We also continue to maintain the LDM Alchemy repository as a collection of LDM processing scripts in Python. Currently this includes the code powering the AWS NEXRAD archive as well as the program that reconstitutes NOAAPORT GOES-16/17 imagery. As we transition more of our internal data processing to Python, this repository will hold those scripts.

Progress has been made on the following:

- Expanding Siphon's scope beyond TDS to a variety of useful atmospheric science datasets, such as the National Data Buoy Center data holdings.
- Siphon 0.8.0 was released, including support for basic HTTP authentication, a client for downloading National Data Buoy Center text data, and built-in support to open datasets using XArray.

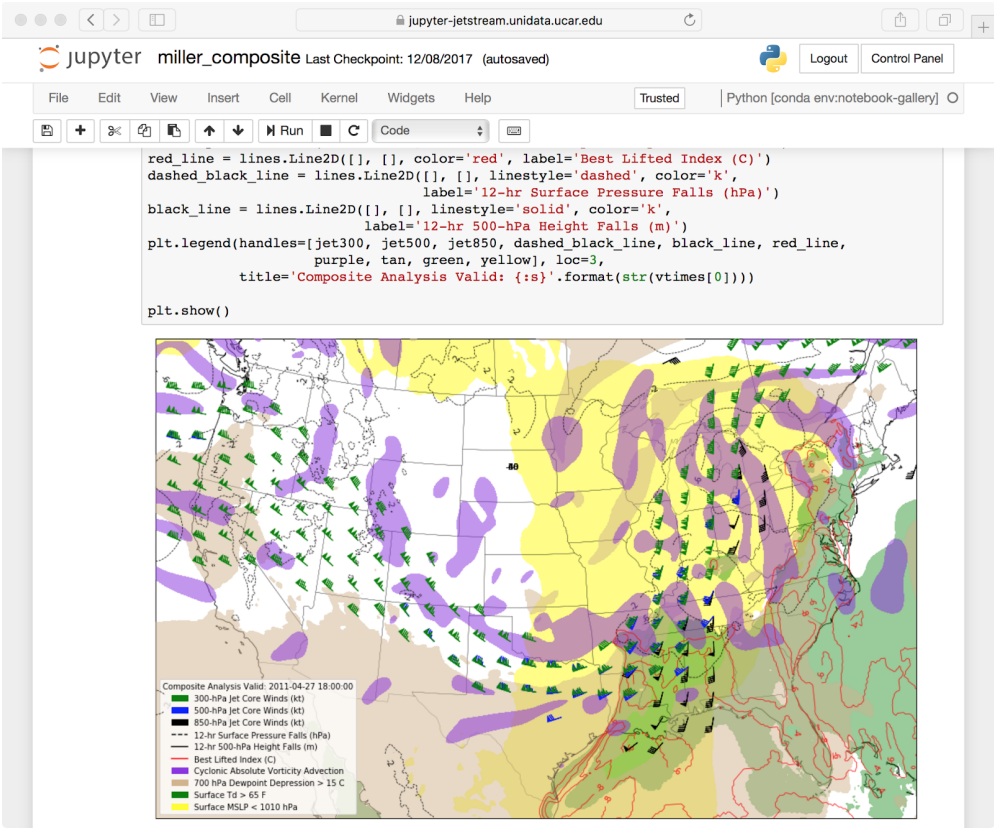
External Participation

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. As participants in a broader Open Source software ecosystem, the Python team regularly encounters issues in other projects relevant to our community's needs. As such, we routinely engage these projects to address challenges and submit fixes. We also continue to host Jeff Whittaker's netCDF4-python project repository; Jeff continues to be the active maintainer of the project. The overall involvement helps ensure that important portions of our community's Python stack remain well-supported.

Progress has been made on the following:

- Ryan May was invited to again present on MetPy at the "Workshop on developing Python frameworks for earth system sciences" hosted by ECMWF in October 2018.
- Ryan May and John Leeman attended SciPy 2018.
- We continue to engage with the Pangeo project, a grass-roots effort to develop a community stack of tools serving the atmospheric, oceanic, land, and climate science, including attending the Pangeo Developer's meeting in August 2018. This engagement is enhanced by work on the Pangeo EarthCube award.
- Ryan May has been engaging with NCAR, especially the NCL team, as they increasingly migrate towards more Python.
- Ryan May has helped fix several issues in CartoPy and continues to help push CartoPy forward. He also continues participating on the matplotlib core development team.
- We also continue to actively engage with the xarray, numpy, and pint projects

JupyterHub Server on Jetstream



JupyterHub is a multi-user server for running Jupyter notebooks and other technologies. We deployed a JupyterHub server with pre-installed Unidata Python notebooks on the NSF-funded Jetstream cloud. The aim here is to provide a zero install solution for running Unidata notebooks for our community in a classroom or workshop setting. After an initial testing phase, we have opened the server to anyone with an XSEDE portal login.

While this demonstration server is a promising start, with, at times, frequent daily use from some users, it will not scale to a large classroom or workshop setting (> ~10 users). We continue researching and employing technologies that allow us to better take advantage of the elasticity of the cloud such as Kubernetes and the Zero to Jupyterhub project. Part of our Jetstream XSEDE allocation includes Extended Collaborative Support Services (ECSS) staffing resources. With those staffing resources, we collaborated with Semir Sarajlic (Georgia Tech) and Suresh Maru (Indiana University) to implement a JupyterHub server using of Kubernetes technology to better take advantage of the horizontal scalability of the cloud. We presented our work at the PEARC '18 conference in Pittsburgh, PA (please see Cloud Status Report).

Recognizing that Unidata time and resources are limited, we have heavily leveraged open-source solutions (e.g., JupyterHub Docker container) for this server and simply deploying, configuring and customizing those open source technologies, resulting minimal in-house development.

Python for AWIPS

We continue to update the Python Data Access Framework (python-awips) package with the latest changes from the AWIPS baseline. This package is used in both AWIPS and GEMPAK for remote retrieval of AWIPS data (grids, geometries, and imagery), as well as independently in Jupyter Notebooks https://nbviewer.jupyter.org/github/Unidata/python-awips/tree/master/examples/notebooks/

The next python-awips release will be 18.1.1 to correspond with the upcoming release of AWIPS EDEX and CAVE (new naming convention from previous 0.9.11, 0.9.12, etc.)

NWS changes to python-awips which are being incorporated into the next release include:

- New support for optional Qpid SSL connection.
- New class **GetGridLatLonRequest** for lat/lon requests to be made separately from gridded data requests, allowing for "lazy loading" of grids where the library waits until the lat/lon for a grid is needed before loading. Examples of when this is True are when the lat/lon information is not used, or when it is used if certain conditions are met. It could be set to False if it is guaranteed that all lat/lon information is needed.
- New setters and getters for geospatial envelope and coordinate reference system (WKT) in **GetGridDataResponse.py**.New **DataQueue** convenience class for using the notifications feature. This is a collection that, once connected to EDEX by calling start(), fills with data as notifications come in. Runs on a separate thread to allow non-blocking data retrieval.
- New data notification class **PyGridNotification** which improves performance over requesting by datatime since it requests only the parameter that the notification was received for (instead of this and all previous parameters for the same forecast hour). However, this still fails for derived requests.
- Apache Thrift updated to version 0.10.0.

Ongoing Activities

We plan to continue the following activities:

- Unidata Python training workshop
- Growing Siphon as a tool for remote data access across a variety of services
- Growing and developing MetPy as a community resource for Python in meteorology
- Continued participation in the scientific Python community as advocates for the atmospheric science community
- Working with JupyterHub as a way to facilitate data-proximate analysis
- MetPy Mondays for engaging the community
- As resources and time permit, continue growing the Online Python Training project by writing Jupyter notebooks specifically targeted towards teaching the geoscience community programming concepts.

New Activities

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

- Work on native Python support for decoding GRIB3 formatted files
- Teach workshop at San Jose St. University
- Attend ECMWF Python workshop in Reading, UK

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

- Teach another short course on MetPy at AMS 2019
- Present annual update on Python libraries at AMS 2019
- Teach another regional MetPy/Siphon workshop as outlined in the MetPy NSF award

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
- Restructure our annual Python workshop to be a full week with introduction to Python/git, intermediate with MetPy/Siphon/etc., and developer hack-day

Relevant Metrics

MetPy

- 98% test coverage
- Watchers: 40
- Downloads for the releases made in the last year (only Conda for now):
    - 0.6.0: 2433
    - 0.6.1: 2156
    - 0.7.0: 4243
    - 0.8.0: 3071
    - 0.9.0: Not released
    - 0.9.1: 570
- Since 1 April 2018
    - Active Issues: 132 (67 created, 67 closed)
    - Active PRs: 102 (72 created, 76 closed)
    - External Issue Activity: 30 opened, 102 comments
    - External PR Activity: 16 opened, 44 comments
    - Unique external contributors: 31
    - Stars: 57 (306 total)
    - Forks: 1 (133 total)
    - Commits: 245
- Since 1 October 2017
    - Active Issues: 264 (184 created, 136 closed)
    - Active PRs: 186 (176 created, 174 closed)
    - External Issue Activity: 57 opened, 234 comments
    - External PR Activity: 47 opened, 118 comments
    - Unique external contributors: 61

    - Stars: 111 (306 total)
    - Forks: 7 (133 total)
    - Commits: 560

Siphon

- 97% test coverage
- Watchers: 12
- Downloads for the last year (only Conda for now):
    - 0.6.0: 2336
    - 0.6.1: 4213
    - 0.7.0: 5384
    - 0.8.0: 1006
- Since 1 April 2018:
    - Active Issues: 29 (16 created, 19 closed)
    - Active PRs: 21 (19 created, 19 closed)
    - External Issue Activity: 9 opened, 20 comments
    - External PR Activity: 5 opened, 5 comments
    - Unique external contributors: 11
    - Stars: 16 (81 total)
    - Forks: 0 (33 total)
    - Commits: 69
- Since 1 October 2017
    - Active Issues: 48 (37 created, 34 closed)
    - Active PRs: 49 (49 created, 47 closed)
    - External Issue Activity: 14 opened, 34 comments
    - External PR Activity: 96 opened, 17 comments
    - Unique external contributors: 19
    - Stars: 30 (81 total)
    - Forks: 2 (33 total)
    - Commits: 193

Python-AWIPS

- Downloads for the last month: 641
- Downloads for 2018: 3,455
- Downloads for the last 12 months: 5,387
- All-time downloads: 14,023

Unidata Python Workshop

- Watchers: 27
- Since 1 April 2018
    - Active Issues: 40 (35 created, 29 closed)
    - Active PRs: 38 (22 created, 24 closed)
    - External Issue Activity: 0 opened, 0 comments
    - External PR Activity: 0 opened, 0 comments
    - Unique external contributors: 0

    - Stars: 13 (88 total)
    - Forks: 0 (65 total)
    - Commits: 61
- Since 1 October 2017
    - Active Issues: 78 (55 created, 56 closed)
    - Active PRs: 57 (51 created, 52 closed)
    - External Issue Activity: 1 opened, 0 comments
    - External PR Activity: 2 opened, 9 comments
    - Unique external contributors: 2
    - Stars: 27 (88 total)
    - Forks: 4 (65 total)
    - Commits: 157

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 accessing remote data, including the THREDDS data server. Moreover, Python technology coupled with the HTML5 Jupyter Notebook technology has the potential to address "very large datasets" problems. Jupyter Notebooks can be co-located to the data source and accessed via a web browser thereby allowing geoscience professionals to analyze data where the data reside without having to move large amounts of information across networks. This concept fits nicely with the "Unidata in the cloud" vision and the goals outlined Unidata 2018 Five-year plan. Lastly, as a general purpose programming language, Python has the capability to analyze and visualize diverse data in one environment through numerous, well-maintained open-source APIs. The additional development of MetPy fills the need for domain-specific analysis and visualization tools in Python.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The TDS catalog crawling capabilities found in siphon will facilitate access to data remotely served by the Unidata TDS, as well as other TDS instances around the world.

4. **Build, support, and advocate for the diverse geoscience community**
   Based on interest from the geoscience community, Unidata, as part of its annual training workshop, now hosts a three day session to explore Python with Unidata technology.

Also, to advance the use of NetCDF in Python, Unidata has promoted Jeff Whitaker's NetCDF4-python project, including hosting its repository under Unidata's GitHub account. Unidata is initiating a project to provide online Python training specifically targeting geoscience students. Unidata is also fostering some community development of meteorology-specific tools under the MetPy project.

Prepared *September 2018*

Status Report: Support

April 2018 - September 2018

Jennifer Oxelson, Tom Yoksas, UPC Staff

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Is the support that we provide sufficient for the community's needs?

Activities Since the Last Status Report

Training

- In 2018 the UPC has been focusing its in-person training efforts on regional workshops
- Additional resources will be directed towards developing online training materials.

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

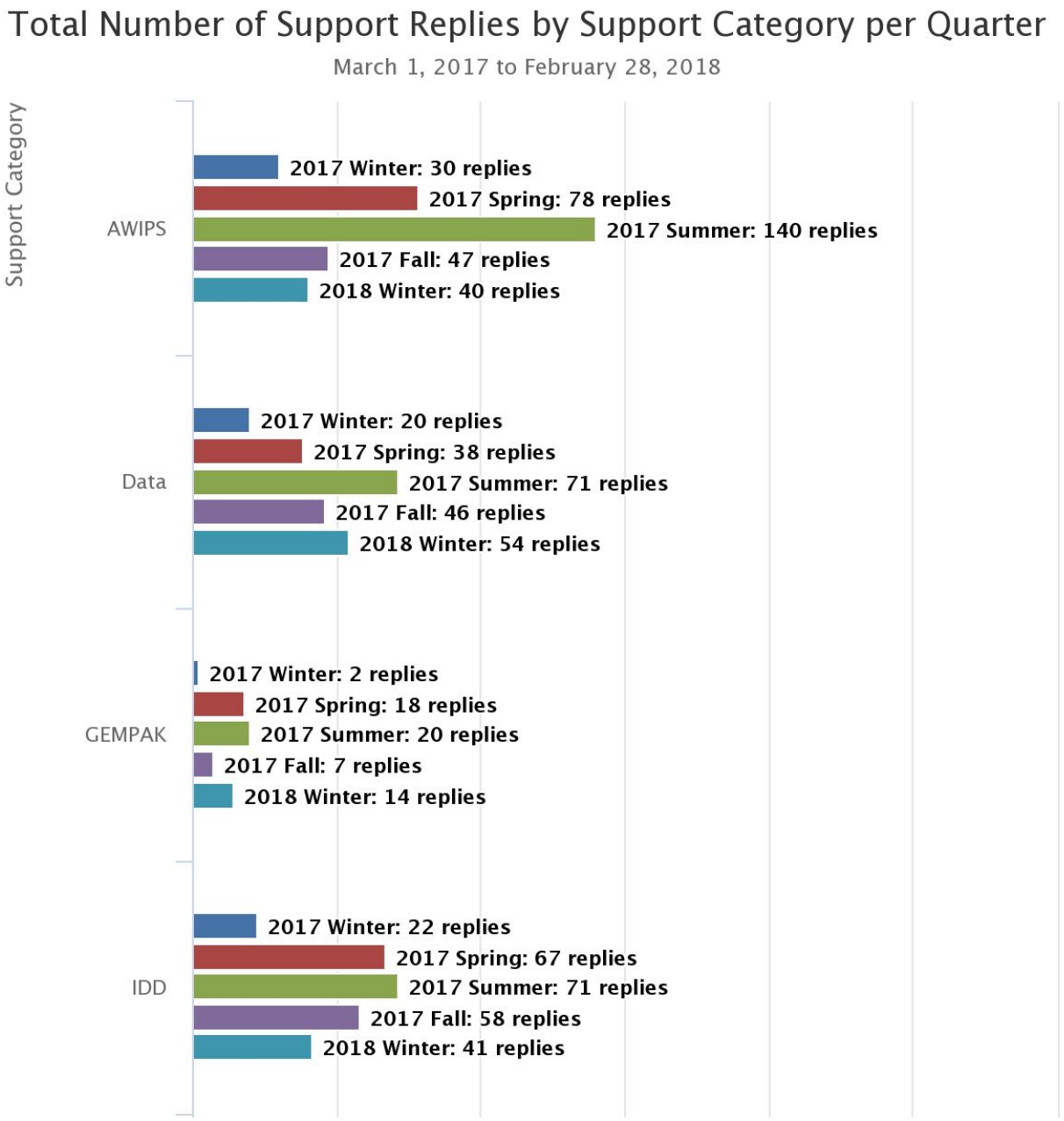
Relevant Metrics

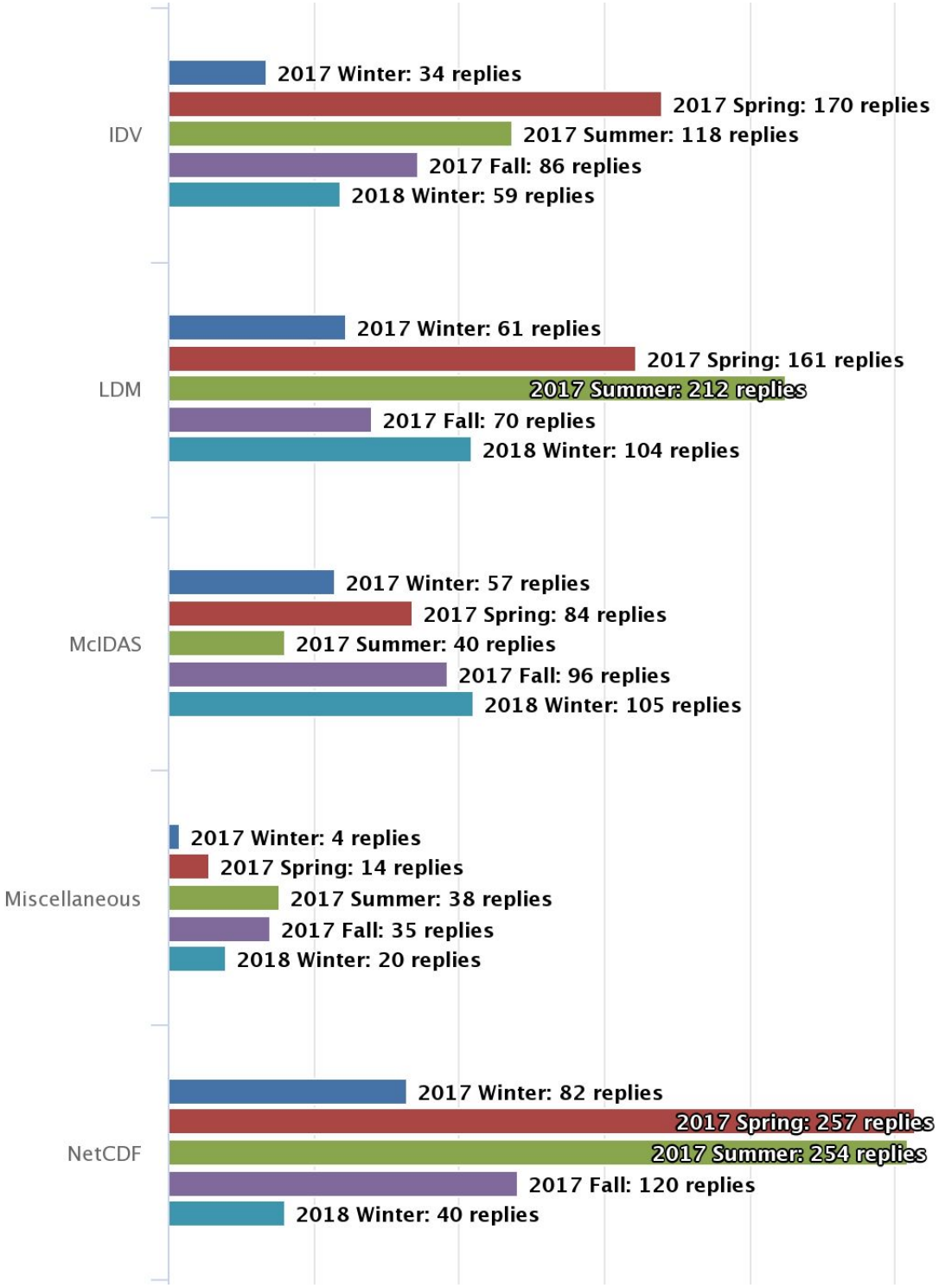
Since January 26, 2006 over 56650 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

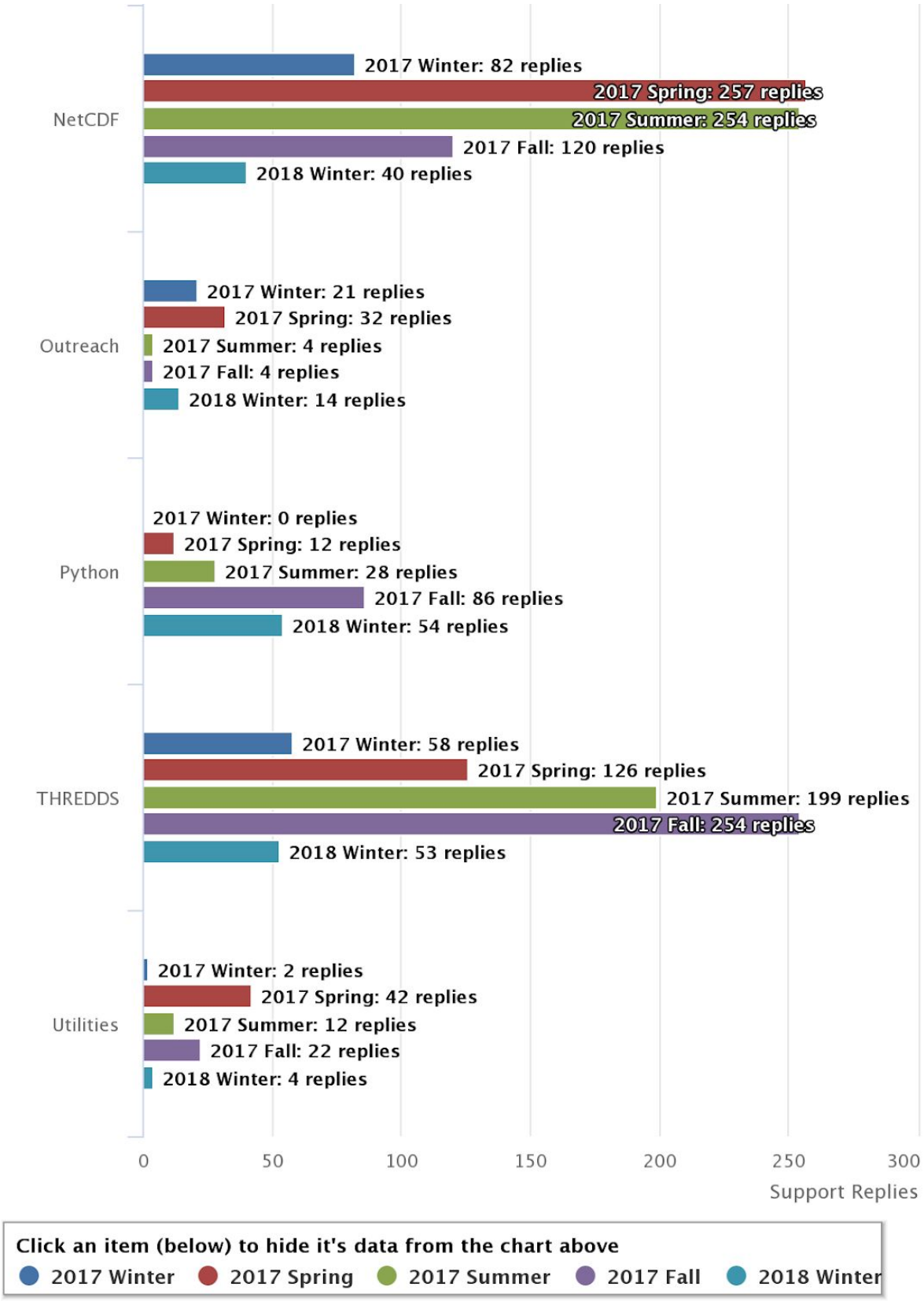
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 14 month period from March 1, 2017 until February 28, 2018. *An updated snapshot that covers the period through August, 2018 will be provided in a future update to this document.*

The quarters shown are defined as: Winter: January, February, March; Spring: April, May, June; Summer: July, August, September; Fall: October, November, December.


Total Number of Support Replies by Support Category per Quarter
March 1, 2017 to February 28, 2018
Support Category
AWIPS
2017 Winter: 30 replies
2017 Spring: 78 replies
2017 Summer: 140 replies
2017 Fall: 47 replies
2018 Winter: 40 replies
Data
2017 Winter: 20 replies
2017 Spring: 38 replies
2017 Summer: 71 replies
2017 Fall: 46 replies
2018 Winter: 54 replies
GEMPAK
2017 Winter: 2 replies
2017 Spring: 18 replies
2017 Summer: 20 replies
2017 Fall: 7 replies
2018 Winter: 14 replies
IDD
2017 Winter: 22 replies
2017 Spring: 67 replies
2017 Summer: 71 replies
2017 Fall: 58 replies
2018 Winter: 41 replies


IDV
2017 Winter: 34 replies
2017 Spring: 170 replies
2017 Summer: 118 replies
2017 Fall: 86 replies
2018 Winter: 59 replies
LDM
2017 Winter: 61 replies
2017 Spring: 161 replies
2017 Summer: 212 replies
2017 Fall: 70 replies
2018 Winter: 104 replies
McIDAS
2017 Winter: 57 replies
2017 Spring: 84 replies
2017 Summer: 40 replies
2017 Fall: 96 replies
2018 Winter: 105 replies
Miscellaneous
2017 Winter: 4 replies
2017 Spring: 14 replies
2017 Summer: 38 replies
2017 Fall: 35 replies
2018 Winter: 20 replies
NetCDF
2017 Winter: 82 replies
2017 Spring: 257 replies
2017 Summer: 254 replies
2017 Fall: 120 replies
2018 Winter: 40 replies


NetCDF
2017 Winter: 82 replies
2017 Spring: 257 replies
2017 Summer: 254 replies
2017 Fall: 120 replies
2018 Winter: 40 replies
Outreach
2017 Winter: 21 replies
2017 Spring: 32 replies
2017 Summer: 4 replies
2017 Fall: 4 replies
2018 Winter: 14 replies
Python
2017 Winter: 0 replies
2017 Spring: 12 replies
2017 Summer: 28 replies
2017 Fall: 86 replies
2018 Winter: 54 replies
THREDDS
2017 Winter: 58 replies
2017 Spring: 126 replies
2017 Summer: 199 replies
2017 Fall: 254 replies
2018 Winter: 53 replies
Utilities
2017 Winter: 2 replies
2017 Spring: 42 replies
2017 Summer: 12 replies
2017 Fall: 22 replies
2018 Winter: 4 replies
0
50
100
150
200
250
300
Support Replies
Click an item (below) to hide it's data from the chart above
2017 Winter
2017 Spring
2017 Summer
2017 Fall
2018 Winter

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

Category	eSupport Departments
AWIPS	Support AWIPS
Data	Support CaseStudy, Support CONDUIT, Support Datastream, Support LEAD, Support Level II, Support NOAAPORT, Support SUOMINET
GEMPAK	Support GEMPAK
IDD	Support IDD, Support IDD Antarctica, Support IDD Brasil, Support IDD Cluster, Support IDD SCOOP, Support IDD TIGGE
IDV	Support IDV, Support IDV Storm, Support McV, Support VisAD
LDM	Support LDM
McIDAS	Support McDevelop, Support McIDAS
Miscellaneous	Administration, Development, Plaza, Staging Folder, Support, Support eSupport, Support Miscellaneous, Support Platforms, Support Plaza, Systems
NetCDF	Support LibCF, Support netCDF
Outreach	Outreach, Polcomm, Support Egrants, Support News, Support Outreach, Support Workshop, Usercomm
Python	Support Python
RAMADDA	Support RAMADDA
THREDDS	Support netCDF Java, Support THREDDS
Utilities	Support LDM-McIDAS, Support netCDF Decoders, Support netCDF Perl, Support OPeNDAP, Support Rosetta, Support UDUNITS

Comments

- The total support provided by the UPC continues to be substantial: yearly totals have been relatively constant for the past several years. Overall support activities vary by somewhat by quarter. Spikes in support for individual packages is largely correlated with the releases of new distributions of the packages supported.
- Support for netCDF continues to be substantial, and is understandable given the very large number of users of the package worldwide.
- Support for the legacy visualization packages GEMPAK and McIDAS has decreased over the past couple of years, most likely due to GEMPAK users investigations of use of AWIPS and Python/MetPy.
- Support for AWIPS has steadily increased and has exceeded that for GEMPAK over the past couple of years.
- Support for Python scripting using MetPy is growing markedly.
- Support for LDM, IDD, and Data continue at a high levels and show some variability throughout the year.

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 Metrics

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**

Unidata User Support enables access to geoscience data by supporting the use of tools created and/or supported by the UPC.

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

Prepared *September 2018*

Status Report: THREDDS

April 2018 - September 2018

Sean Arms, Ethan Davis, Dennis Heimbigner, Ryan May

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. We are hiring! If you know someone who would love to work on THREDDS related projects at Unidata, send them our way!

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. An update regarding cloud efforts related to the TDS, including the popular Docker container effort, can be found in the Cloud Computing Activities Status Report.

Released netCDF-Java / TDS version 4.6.11 (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. “Maintenance mode”, which includes user support and bug fixes, continues to take up quite a bit of resources.
- The THREDDS team now conducts automated security scans on our dependencies to ensure that we are not using external libraries with open vulnerabilities.

Focus on netCDF-Java / TDS (Beta) v5

We are now on beta4 of the TDS.

Progress has been made on the following:

- Thanks to our Summer Inter, Hailey Johnson:
    - a templating system has been added to the TDS, allowing for easy customization of the HTML interface
    - a new viewer has been added which generates Jupyter Notebooks - these can be customized on a per-dataset level.
- The Nexus Repository Manager at https://artifacts.unidata.ucar.edu has been

upgraded from version 2 to version 3 and it will now host **all** build artifacts. For users, this means:
    - Developers should update their software builds to pull netCDF-Java/TDS artifacts from the Maven repository at https://artifacts.unidata.ucar.edu/repository/unidata-all/.
    - Documentation for versions 5 and later of netCDF-Java/TDS will reside at https://artifacts.unidata.ucar.edu/repository/thredds-doc/ and—more conveniently—at https://docs.unidata.ucar.edu/thredds (specifically, https://docs.unidata.ucar.edu/thredds/5.0.0-SNAPSHOT/userguide/index.html)
- The configuration management tool Ansible has shown great promise as a way for users to be able to deploy TDS and other Unidata software in an automated fashion.
    - Given Christian's departure, we are struggling to keep the ansible build working as it is such a radical departure from the norm.
- DAP4 in the TDS has been updated to be consistent with the specification and to successfully allow the netCDF-C DAP4 and NetCDF-Java libraries to read DAP4 responses from the TDS.
- New Coverage data type allows for subsetting across array boundaries (often called the "seam" problem).
- Uses the new edal-java based ncWMS 2.0 server, as well as javascript client Godiva3.
- CatalogScan feature allows for incremental updating of TDS catalogs without the need to restart Tomcat.
- Upload/Download support has been added to TDS. This now includes an upload web form accessible as http://.../thredds/upload.
- Unit and Integration tests are passing in 5.0. This is a big step towards releasing a beta.
- ncSOS has been integrated into the TDS distribution (as part of the OIIP project—see the Rosetta section for more details)
- Access to the netCDF-C library via JNI is now thread-safe so that the HDF5 library no longer needs to be built with thread-safe support.
- The license for netCDF-Java and the TDS has been updated to a BSD-3 clause licence. See https://www.unidata.ucar.edu/blogs/developer/entry/thredds-licence-change for more information.

Dependencies, challenges, problems, and risks include:

- The loss of Christian has been extremely difficult, as he was the only staff member dedicated to netCDF-Java / TDS work.
- Maintenance of the 4.6.x line of netCDF-Java and TDS continues to have a large impact. The goal of beta testing TDS 5 is to ensure that the current capabilities of 4.6.x are working in the new version (and if some bugs get fixed in the process, even better!). Beta testing by our users will be critical, and so far we have had several community members offer their help (special thanks to Rich Signell!).

Rosetta

Rosetta continues to progress thanks to support from a NASA ACCESS grant (the Oceanographic In-situ data Interoperability Project, or **OIIP**), in which Unidata is partnering

with the PO.DAAC at JPL and UMASS-Boston. The project ends this month.

Progress has been made on the following:

- Support for the NCEI NODC netCDF v2.0 templates (metadata standards)
- Extension of the NCEI templates to support metadata critical to the use of electronic tagging datasets
- Support automated transformation of output from electronic animal tagging datasets in the Electronic Tag Unified File Format (eTUFF) format via Rosetta.
- Working to create a unified workflow for the gui wizard interface that allows for selection of which metadata standards to use when determining recommended/required metadata
- Engaging with the netCDF Linked Data initiative to define best practices identifying netCDF metadata to a particular metadata standard.
- The entire front end has been reworked, and no longer depends on jWizard. Thus, the outdated jQuery lib is not longer an issue.

Dependencies, challenges, problems, and risks include:

- As with many other projects, lack of resources puts Rosetta at risk. External funds help, but rarely provide the ability to bring on new staff members, which results in taking resources away from other projects.

Ongoing Activities

We plan to continue the following activities:

- Documentation updates - We are reworking the tutorial material for TDS v5.0 with the goal of enabling asynchronous training. The material will undergo a major overhaul to include the use of Docker containers, video snippets, and other new forms of training tools. The first pass at the overhaul is now complete.
- Maintain thredds.ucar.edu and keep up with the addition of new datasets to the IDD.
    - GOES-17 data, with tiles stitched together using python, is available on our test TDS.
- Continue development of the TDS python client siphon, as well as extend its functionality to interface with other web services and servers.

The following active proposals directly involve THREDDS work:

- Thanks to Rich Signell, we, along with Axiom Data Science, submitted and were **awarded** a NOAA IOOS grant. The proposal was entitled “A Unified Framework for IOOS Model Data Access”, and the goal to enable support of the UGRID specification within the THREDDS stack, as well as create a GRID featureType to allow for serving large collections of gridded datasets (including UGRID). This work will fund a Undiata staff member at 0.5 FTE for two years, as well as two co-PIs at Axiom Data Science at a slightly lower level.
- The NASA ACCESS award with JPL is ending this month. The award is titled:

"Leveraging available Technologies for Improved interoperability and visualization of Remote Sensing and in-situ Oceanographic Data at the PO.DAAC" and was submitted with JPL/PO.DAAC. [Rosetta]. More information can be found at https://oiip.jpl.nasa.gov/
- 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. Currently in a no-cost extension period.

New Activities

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

- Hiring a new developer to work on netCDF-Java / TDS
- Officially advertising a public TDS 5.0 Test Server [currently found at http://thredds-test.unidata.ucar.edu/thredds/catalog.html]
- Getting TDS v5.0 to a stable release
- Ryan May and Sean Arms are officially involved with the GRIB-3 effort at the WMO. Work is being done to create native java and python decodes for the new version as independent implementations to validate the GRIB-3 specification.

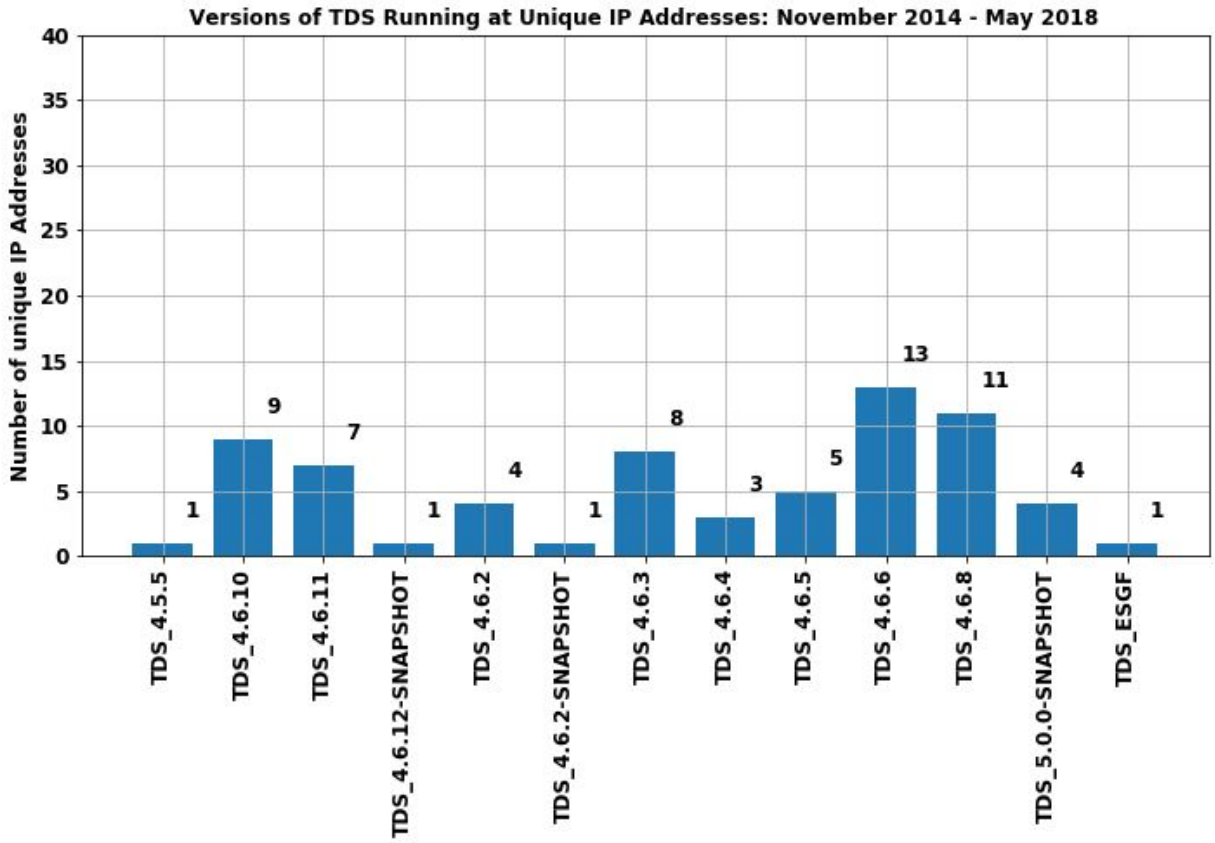
Over the next twelve months, we plan to organize or take part in the following:

- Supporting Zarr and, potentially, TileDB at the IOSP layer of the CDM.
- Support the UGRID spec within the CDM
- Support the netCDF-CF simple-geometry spec within the CDM (work under way by staff members at USGS funded by a Water Resources Research Institute)
- Create a TDS plugin layer for external services (potentially use Project Jigsaw from Java 9)
    - 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

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

- NcML and FMRC collections have many problematic warts, including unreliable caching. Tackling these issues will be critical once netCDF-Java and TDS 5 are released.
- The scalability of a serverless architecture is very attractive, but would require a re-architecture of the TDS, likely to something resembling a microservice architecture. While the time horizon on this kind of transition is long, we plan on exploring options with some TDS capabilities, such as the catalog service or the netCDF Subset Service.

Relevant Metrics



12,061 unique IPs started up thredds from November 2014 through May 2018, **68** of which are publicly accessible servers. “Publically accessible” is defined as the following URL patterns being accessible with an HTTP GET requests with a return status less than 400 as well as content that contains xml:

http(s)://<ip address>/thredds/catalog.xml
http(s)//<ip address>:8080(8443)/thredds/catalog.xml
http(s)://<ip address>/thredds/catalog/catalog.xml
http(s)//<ip address>:8080(8443)/thredds/catalog/catalog.xml

You may notice that the number of publicly accessible TDSs decreased just over half since our last report (now at 68). This is due to a new check that, in addition to being resolvable, the server response is actually an xml file. Many of the of the previously counted “publicly accessible” TDSs that are now excluded are AWS 404 html pages (and in some case, not so “PG” ad pages).

This information is only known for servers running v4.5.3 and above. There are many reasons why these number are so different. The differences could be due to:

- People testing the TDS on their local machine, but not actually running a server (most likely the cause for the majority of the difference)
- A TDS running behind a proxy server may not be “seen” in this analysis as publicly reachable at the tested url pattern (<server>/thredds/catalog.xml). For example, a TDS running behind a proxy might be configured to respond to mytds.<server>/catalog.xml, and so our check for mytds.<server>/thredds/catalog.xml would not work..
- The TDS server may be running behind a firewall that does not allow public access.
- A TDS running in the past is no longer running today.

Note 1: the vast majority of the publicly accessible servers are running v4.6.3 or above (v4.6.11 was the most current release during this period, and was released on 11 December 2017).

Note 2: there are some odd looking versions of the TDS being reported in the log files, such as TDS_4.28.x. It is likely these version numbers are actually generated by software that is being built on top of the TDS or applications that bundle the TDS as part of a deployment package (perhaps, given the decrease in the number, these are ESGF nodes?).

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 (e.g. ERDDAP, Panoply, VERDI, etc) and on the TDS (ESGF, LAS, ncWMS, MyOcean, etc).

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The Common Data Model (CDM) / netCDF-Java library is one of the few general-purpose implementations of the CF (Climate and Forecast) metadata

standards. Current active efforts in CF that we are involved with include use of the extended netCDF-4 data model (CF 2.0) and for point data (Discrete Sampling Geometry CF-DSG).

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

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

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

Prepared *August 2018*