Strategic Advisory Committee Agenda

PSU, Walker Building, Room 529

Tuesday, April 16, 2019

8:30 - 9:00 Doors Open

09:00 - 9:30 Administrative Items (Sepi Yalda/Staff)

- Review of Action Items
- Date of Fall meeting (Users Committee suggestion: Joint meeting held week of September 30 - October 4)
- Update from Users Committee meeting

09:30 - 10:30 Director's Report (Mohan Ramamurthy)

10:30 - 10:45 Break

10:45 - 11:30 Budget Report

11:30 - noon NSF Update (Bernard Grant)(Time to be confirmed)

Noon - 1:15 Lunch - Café Laura

1:15 - 2:15 Business Development Report (Josh Young)

- Deployment of AWIPS
- Evaluation of data delivery
- University partnerships (proposed approach)

2:15 - 2:50 Opportunities and approaches to engage underrepresented populations (Mohan/Josh)

2:50 - 3:00 Break

3:00 - 3:30 Penn State Meteorology Weather Station and Visualization Tour (Chuck)

3:30 - 4:30 Unidata Products and Services used at Penn State (Chuck/speaker TBC)

4:30 - Adjourn

6:30 - Discussion of the day's proceedings over dinner at The Federal Taphouse (130 S Fraser Street, State College) (map)

Wednesday, April 17, 2019

8:30 - 9:00 Doors Open

9:00 - 10:00 Partnerships around education and workforce development (General Discussion)

10:00 - 10:15 Break/Group Photograph

10:15 - 10:45 NASA Update (Chris Lynnes)

10:45 - 11:30 Restructure of technical teams and implementation of new core award (Ethan/Ryan)
11:30 - Noon Actions/AOB
Noon Adjourn

Status Report: AWIPS and GEMPAK

September 2018 - March 2019

Michael James

Activities Since the Last Status Report

AWIPS

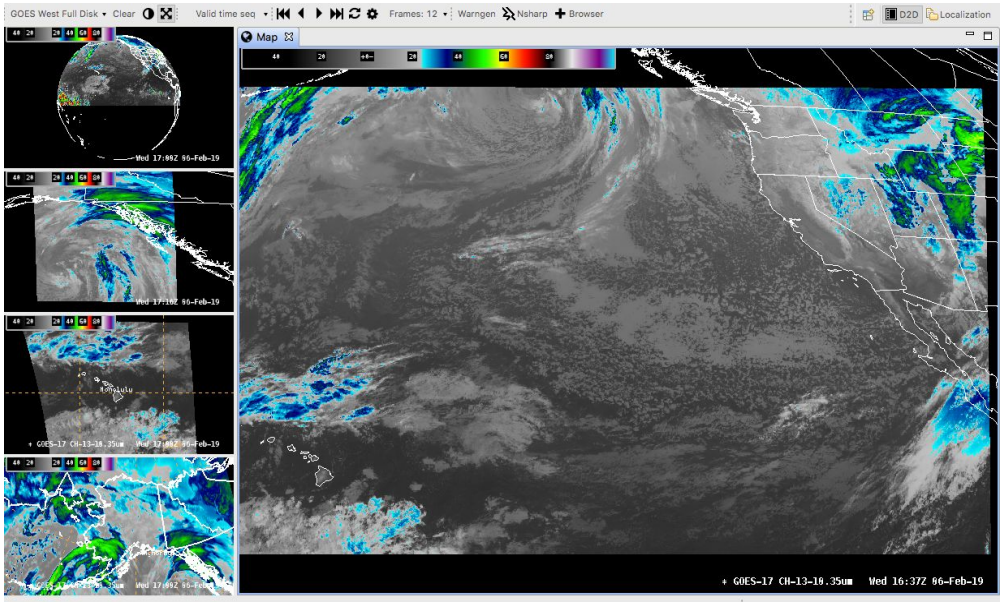
Unidata's Jetstream EDEX server continues to serve real-time AWIPS data to CAVE clients and through the python-awips data access framework (API). By offloading the processing of one very large data feed (NEXRAD3) to a separate EDEX Ingest Node, the current implementation of **edex-cloud** is now cabable 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.

Software Releases

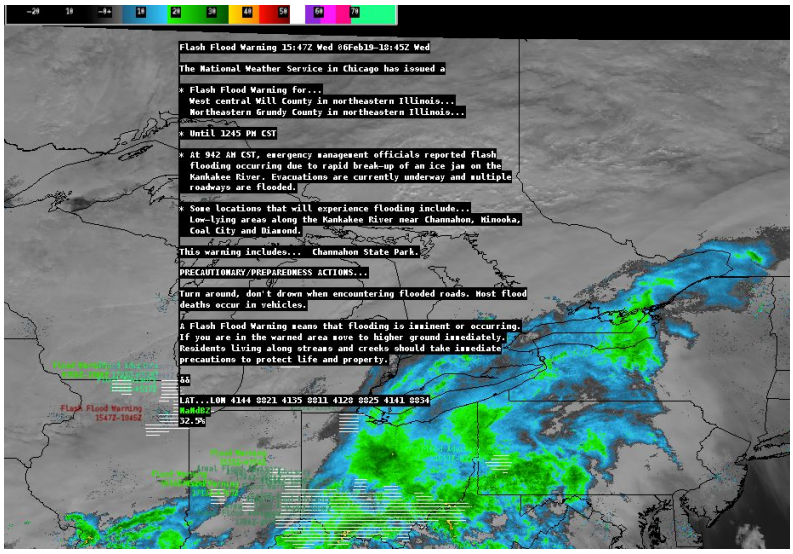
18.1.1-5 Feb 2019

AWIPS 18.1.1-5 is a major update to both CAVE and EDEX, including a security update to the latest httpd/mod_wsgi packages.

- EDEX localization levels have changed, and Unidata AWIPS no longer supports REGION, DESK, or WORKSTATION levels, instead using only Base, SITE, and USER.
- EDEX administrators should be aware that USER-level localization names have been merged with the WORKSTATION name, that is, what used to be USER *awips* will now be USER *awips@<workstation name>*.
- New GOES Geostationary perspective multi-view bundles loaded from the Satellite menu.



- New menu items for GOES 17 (GOES West) sectors WCONUS, WFD, WMESO-1, WMESO-2, AKREGI, HIREGI.
- WWA (Watch, Warn, Advisory) plugin display updates with phensig color table and full text on sampling. These new bundles are located in the Surface menu, available for flood, wind, marine, as well as *All Warnings* and *Convective Watches*.



- Fix for the Linux GFE Perspective to sync required files to run gfeConfig correctly.
- Removal of REGION, DESK, and WORKSTATION localization level. Any

configuration or setting which used the DESK or WORKSTATION levels will now use SITE instead. This makes for only 3 localization levels to manage in Unidata AWIPS from 18.1.1-5 onward: 1) BASE (default), 2) SITE (e.g. OAX), and 3) USER.
- More efficient thread control for the ingestGrib JVM in `/awips2/edex/etc/ingestGrib.sh` rather than using definitions in `/awips2/edex/conf/resources/`.
- Updated the `/etc/init.d/edex_ldm` wrapper to allow for the use of `ldmadmin clean` as root.
- Fixed a bug with Python environments when running `edex qpid` as root.
- Updated `pqact.conf` to include GOES 17 products, as well as a fix to ingest both METOP-A and B ascat files.

18.1.1-4 (development release) Dec 2018

18.1.1-3 Nov 2018

- Added a new GEMPAK display plugin to the D2D perspective
- Cleaned up paths available to CAVE in the Localization Perspective
- Cleanup grid name aliasing to avoid multiple names for a single grid (NAM40 vs NAMWX)
- Fix to reset the manual ingest endpoint to `/awips2/data_store/ingest/`
- Cleanup/remove GFE service backup utility
- Updates to "edex start" service manager
- Minor updates and fixes merged from NWS development baseline (Hydro apps trace precipitation display, GFE Fosberg fire weather index, GFE Haines index, Grib1 decoding updates)

18.1.1-2 Sep 2018

- “edex status” will mute EDEXregistry if Data Delivery/OGC plugins are not installed. Data Delivery/OGC plugins are not being used currently, but work is ongoing to incorporate them into EDEX (more on this in the future).
- New awips2-python-scipy RPM for CentOS 7.
- FNMOC grid names corrected in awips2-edex-dataplugins.
- Updated awips2-ldm and awips2-edex to remove the edex setup command from post-RPM scripts, and from EDEX startup, to avoid overwriting user edits to the LDM registry.
- "edex setup" will be a required command after install and before starting EDEX services.

18.1.1-1 Sep 2018

- Windows CAVE

  https://www.unidata.ucar.edu/downloads/awips2/awips-cave.msi

  - Requires Python 3.7, Jep (3.8.2), and NumPy (>1.7).
  - `$PYTHONHOME` must be defined for full functionality using gridslice and jep.
  - The MSI installer will attempt to copy the gridslice shared library to `$PYTHONHOME/Dlls`. If the `$PYTHONHOME` environmental variable is not defined, gridslice will not be installed.
  - You can always rerun the installer after defining `$PYTHONHOME` and then check that the file gridslice.pyd is installed in `$PYTHONHOME/Dlls`.
  - Windows CAVE will run without gridslice, but certain bundles which use derived parameters, such as isentropic analyses, will not load.
- New D2D CAVE macOS client
  https://www.unidata.ucar.edu/downloads/awips2/awips-cave-18.1.1-1.dmg
- New CentOS 6 and 7 client

```
wget https://www.unidata.ucar.edu/software/awips2/awips_install.sh
chmod 755 awips_install.sh
./awips_install.sh --cave
```

- New CloudAWIPS release 18.1.1-1
  https://github.com/Unidata/CloudAWIPS/blob/master/README.md
- GOES-17 ABI sectors TCONUS and TFD
- Fix CAVE crash when starting GFE for non-activated site.
- Updates to allow dynamic contributions to the Volume Browser.
- Search By Name has been added to the Localization Perspective.
- Fix serialization errors which prevented IPv6 addresses from connecting.
- Derived Parameter functions synced to client CAVE sessions are now Python 2 and 3-compliant.
- Support for variable header length in PNG-compressed McIDAS AREA files (UNIWISC)
- Database reconfiguration (moved from /awips2/data to /awips2/database/data, now with 'awipsadmin' ownership).
- New roles/permissions framework to replace nwsauth from versions 17.1.1 and before.
- pqact.conf added Jason-3 altimeter data feed
- pqact.conf cleanup of removed products and duplicate pattern actions
- Updates to the Qpid Broker and Localization REST configuration including Protected File status.
- Refactored gridslice.c source code for a Python 3 environment.
- New viz method IGLMesh for easier extensibility
- Ability to handle derived parameter cubes at a single point
- ADE updates for Eclipse 4.6.1
- Geotools updated to 16.4
- Cython updated to 0.28.3
- httpd-pypies updated to 2.4.27

GEMPAK/NAWIPS

GEMPAK release 7.5.1 February 2019

https://github.com/Unidata/gempak/releases/tag/7.5.1

- Fixed build of program dctama
- Supports GOES-16/17 ABI imagery in NetCDF 4 format

  GOES16 and GOES17 ABI products are distributed in NetCDF4 formatted files, and GEMPAK as of release 7.5.1 can now read and display these products in their native formats (it is no longer needed to convert netCDF to McIDAS area format to display in GEMPAK).

  For Full Disk and CONUS sectors, GOES16 and GOES17 ABI products are sent across the Satellite Broadcast Network as tiles, each contributing to a "mosaic" of a full sector scan. The Unidata Program Center has provided an alternate data feed in which all tiles are "stitched" together (thanks to the program ldm-alchemy).

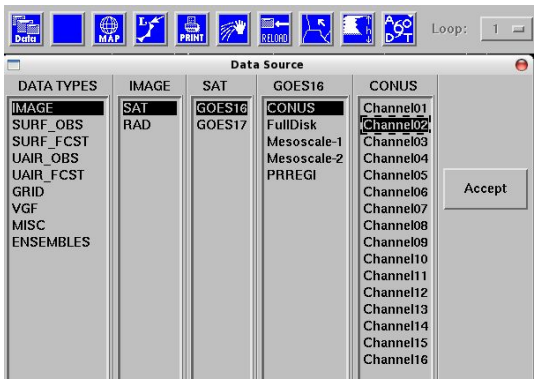
  To request these stitched images, add the following request line to your ldmd.conf:

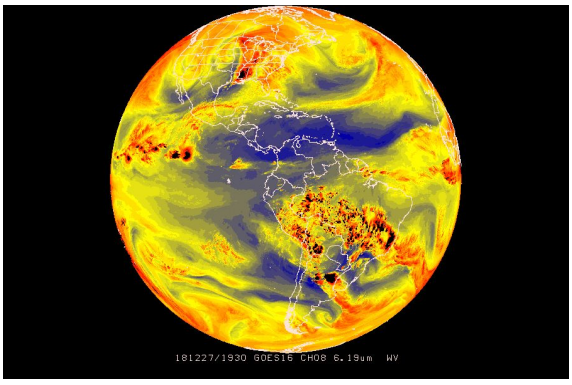
  ```
  REQUEST NIMAGE "GOES" idd.unidata.ucar.edu
  ```

  and the following to your GEMPAK pqact.conf:

  ```
  NIMAGE     .*GOES/(.*)/([^/]*)/([^/]*)/([0-9]{8})/GOES(16|17)_([^/]*)_([0-9]{8})_([0-9]{4})(..)_([^/]*).nc4
         FILE    -close  data/gempak/images/sat/\1/\2/\3/\3_\7_\8
  ```

  This will write to directories in `$SAT` which are selectable in NMAP2 > New Data Source > Image:





GEMPAK release 7.4.5 December 2018

https://github.com/Unidata/gempak/releases/tag/7.4.5

- GEMPAK Python classes refactored for Python 3-compliance and best practice coding convention.
- AWIPS grid naming updates.

GEMPAK release 7.4.3 September 2018

https://github.com/Unidata/gempak/releases/tag/7.4.3

- Full support for GOES ABI fixed grid projection McIDAS AREA format (UNIWISC).
- New AWIPS lutfils for GOES 16 IR and WV channels (https://www.unidata.ucar.edu/staff/mjames/gempak/goes16/).
- Corrected a bug in 7.4.2 snow squall warnings which caused dcwarn to generate false FFW warnings.
- Library updates for future netcdf4 support (zlib 1.2.6 -> 1.2.8, HDF5 1.8.0 -> 1.8.15, netCDF 4.1.3 -> 4.3.3.1).

Activities Ongoing/In-Progress

EDEX and WRF

Using the Jetstream cloud, I have created an EDEX ingest with self-ingested scheduled UEMS (WRF) runs.

Unidata AWIPS release 19.1.1 (release TBD)

- Developing a set of plugins for the ingest and display of USGS river gauge RDB files as time-series pointdata.
- Adding world cities, administrative districts, roads, and other international shapefile imports.
- Menu updates for Watch, Warn, Hydro/QPF, Upper Air products.
- Support for the North American Real-time Ocean Forecast System (RTOFS) grids.
- LDM 6.13.7.
- OGC and Data Delivery plugins have been leveraged to read NOMADS product catalog from a CAVE client. Further work is required to enable user-directed ingest/subscription to specific products.

Future Activities

- Test AWS API for public data set retrieval via EDEX.
- 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 executable for CAVE for wider Linux distro support.
- Add networking to the CAVE Archive Case Study creation tool.

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 *March 2019*

Status Report: *Cloud Computing Activities*

September 2018 - March 2019

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

Unidata Science Gateway on Jetstream

We are continuing to enhance the Unidata Science Gateway on the NSF Jetstream Cloud: http://science-gateway.unidata.ucar.edu in the following broad topic areas:

A JupyterHub for the Unidata Community

In collaboration with the eXtreme Science and Engineering Discovery Environment (XSEDE) Extended Collaborative Support Services (ECSS) team and Indiana University, Unidata has deployed a new JupyterHub server on the NSF Jetstream cloud (https://jupyterhub.unidata.ucar.edu). Contrary to previous attempts in this area, this JupyterHub can accommodate many users because of its use of the Zero to JupyterHub project which leverages technologies like Kubernetes to scale across the cloud computing data center.

Using resources provided by the National Science Foundation's Jetstream cloud-computing platform, Unidata's JupyterHub server is intended to serve as a demonstration of a notebook-based workflow for geoscience activities. After preliminary testing in a variety of situations including workshop and classroom use, Unidata staff are looking to expand beta-testing of the server to the wider community.

Researchers, educators, and students can connect to the server to use a JupyterHub instance that has been pre-populated with Unidata's existing Python notebook collections and the environments to run them. Available projects are

- The Unidata Python Workshop which teaches the use of Python in the atmospheric and oceanic sciences.
- The Unidata Python Gallery, which is a collection of resources using Python for various meteorological tasks.
- Unidata's Online Python Training is a collection of online training materials focused on the use of Python in the atmospheric sciences.

The Unidata JupyterHub has also been configured with the following technologies:

- **JupyterLab**: Unidata's JupyterHub server is equipped with JupyterLab, the next generation web user interface for Jupyter.
- **Terminal**: Users can interact with the JupyterHub environment using a Unix terminal (command line interface). This facility allows users to import resources into the session using tools such as `git` version control or data-fetching tools such as `curl` or `wget`.
- **nbgitpuller**: The nbgitpuller package ensures that any local git repositories remain in sync with their respective canonical repositories. nbgitpuller is loaded to ensure that the prepopulated notebook projects described previously remain up-to-date.

A JupyterHub for Your Classroom or Workshop

In addition to making the demonstration server available for testing by community members, Unidata has deployed several JupyterHub instances for use by workshops and semester-long classes. If you're interested in having Unidata configure a JupyterHub instance for your classroom or workshop, please contact us.

JupyterHub Servers Deployed on Jetstream for Academic Institutions

For the fall 2018 and Spring 2019 Semesters we supported Jupyterhub at the following institutions:

- Tom Narock | Notre Dame of Maryland University
- Keith Maull (UCP), Xin Yang | Southern Arkansas University

Metrics

This JupyterHub server has been experiencing user growth with 22 new users since it was launched a three weeks ago.

Nexrad AWS THREDDS Server Moves from Amazon to Jetstream Cloud

In 2015, as part of the NOAA Big Data Project, Unidata deployed a THREDDS data server on the AWS cloud serving Nexrad data from Amazon S3. Recently, we migrated the Nexrad AWS THREDDS Data server from the AWS cloud onto Jetstream. The previous version of this server had been running a version of the TDS that was out of date. We modernized this server by making use of the THREDDS Docker container. This switch has also yielded cost savings as

Unidata no longer has to pay for data egress out of AWS. We are also able to take advantage of Jetstream's fast Internet2 capability for serving these data. Finally, use of this server was previously restricted to the university community to limit AWS egress charges. With the move to Jetstream, access to these data is now unrestricted.

IDV on the Jetstream Cloud

For clarification this project is not to be confused with CloudIDV. For those who have a Jetstream allocation (e.g., startup, education, research), an IDV image is available when selecting VMs through the Jetstream Atmosphere dashboard. If you do not have a Jetstream allocation, it is possible to obtain a Jetstream Trial Access account. When searching for images, search for "IDV". You will find the Unidata IDV <version> image. Launching that image will then allow you to "Open Web Desktop" where the IDV will be available pre-loaded with the Jetstream plugin. The catalogs available via the IDV Dashboard, Data Choosers will default to the TDS and RAMADDA servers running on Jetstream enabling data-proximate performance, especially for large datasets. You can also select image sizes that are appropriate to your goals. For example, if working with a data intensive bundle, you could choose an m1.medium VM with 6 vCPUS, 16 GB RAM.

Dependencies, challenges, problems, and risks include:

- We are leveraging the Zero To JupyterHub project for the deployment of this JupyterHub. The success of the Unidata JupyterHub depends on the strength of that project which currently has a vibrant open source community.
- Unidata technical staff continue to work closely with Unidata system administrators to ensure cloud VMs (especially on Jetstream) are adhering to Unidata security standards.

AWIPS EDEX in Jetstream Cloud

Unidata continues to provide 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 scale EDEX in the cloud to account for the desired data feed (and size). 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 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, and AWIPS. 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:

A number of these Unidata Docker containers have been updated since the last status report:

- **Tomcat Docker Container** (which the THREDDS and RAMADDA containers depend on) has been security enhanced and has had some bug fixes
- **LDM Docker Container** was released with the 6.13.7 version of the LDM.
- **TDS Docker Container** was released with the 4.6.13 version of the TDS.

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://jupyterhub.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 S3 and the NSF Jetstream cloud. LDM and THREDDS Data Server (TDS) THREDDS Docker software are being employed to deliver these data.
- TDS on Jetstream for level II NEXRAD: 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

- 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

Exploring the Google Cloud Platform with grant from Google

- Creating a “real time” feed of level II RADAR data
- Sandbox to explore the GCP and how it plays with Unidata software stack

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

Unidata Science Gateway

Forthcoming Presentations

- EGU Annual Meeting 2019 | April 7–12, 2019 - Vienna, Austria
- ESIP Summer Meeting 2019 | July 16-19, 2019 - Tacoma, WA
- Gateways 2019 | September 23–25 - San Diego, CA

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

Unidata Science Gateway and JupyterHub

We would like to promote and advertise the science gateway (http://science-gateway.unidata.ucar.edu/) to our community. Specifically, we would like to gain users for our flagship Jupyterhub server: https://jupyterhub.unidata.ucar.edu.

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 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 *March 2019*

Status Report: Community Services

September 2018 - March 2019

Doug Dirks, Jeff Weber, Joshua Young

Areas for Committee Feedback

We are requesting your feedback on the following topics:

All questions are featured as agenda 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:

- Wanted: Student Representative for Unidata Users Committee
- Python-Focused Software Training Workshop at Valparaiso University
- A JupyterHub for the Unidata Community
- Unidata Program Center Welcomes Howard Van Dam
- Python-Focused Software Training Workshop at Metro State University of Denver
- AMS 2019 Conference Highlights from the Unidata Staff
- Python in Action at Valparaiso University
- Unidata Welcomes New Committee Members
- Successful Python-Focused Software Training Workshop at Jackson State University
- Steven Lazarus Receives 2018 DeSouza Award
- 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. We're starting to make progress working with committee members, but there is more to do.

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:

- Planning for participation in at the Meteorological World Technology Expo and OceanObs19
- 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 follow-up (date and topic TBD)
- 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
- Active participation in CUAHSI HIS (Hydrologic Information System)
- Artive participation in the Hydroshare Advisory Committee

New Activities

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

- Expanded emphasis on cloud-related activities
- Development of a Users Committee workshop follow-on event

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:

- 54,596 unique pageviews (60,695 in previous period)
- 10.3% of total unique pageviews (11.3% in previous period)

Top community pages

1. All blog pages
   42529 unique pageviews (43347 in previous period)
   78% of total community pageviews (71% in previous period)
2. www.unidata.ucar.edu/events
   4884 unique pageviews (11641 in previous period)
   9% of total community pageviews (19% in previous period)
3. www.unidata.ucar.edu/community
   4321 unique pageviews (2832 in previous period)
   8% of total community pageviews (5% in previous period)
4. www.unidata.ucar.edu/about
   2406 unique pageviews (2407 in previous period)
   4% of total community pageviews (4% in previous period)

Social media statistics, March 4, 2019

1. # of Twitter followers: 1016 (up from 904)
2. # of Facebook followers: 788 (up from 754)

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

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

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

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

Status Report: Community Equipment Awards

Sponsored by the National Science Foundation

September 2018 - March 2019

Admin Group

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Possible theme(s) for the 2020 Unidata Community Equipment Awards;
2. Please consider volunteering to serve on the 2020 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 2024 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 last year's 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 November 16, 2018 with a March 08, 2019 submission deadline. The Review Panel will meet on March 22 after the Users Committee Meeting at the Unidata Program Center to make recommendations for proposals to be funded.

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.

Status Report: GOES-R Class Satellites

September 2018 - March 2019

Mike Schmidt, Tom Yoksas

Questions for Committee Members

- A reorganization of the content of the LDM/IDD NIMAGE and UNIWISC datastreams is in progress; pleases see the description below for detail

  Given the reorganization, what image coverages and spatial and temporal resolutions should be considered for the NIMAGE and/or UNIWISC IDD feeds? For instance, should coverages formerly provided by the NWS via NOAAPort be added to the IDD if feasible?

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

  We currently provide access via the IDD (push), THREDDS Data Server (pull), McIDAS ADDE (pull) and AWIPS EDEX (pull).

- Other questions?

Activities Since the Last Status Report

- Added distribution of GOES-17 GRB products to the same IDD feed (SATELLITE) that has been providing GOES-16 GRB products for the past year.

  GOES-17 ingest is being done on a 4.5 m satellite dish that was installed at the NCAR Mesa Lab using funds provided by NOAA's GOES-R office. GOES-16 ingest in UCAR is being done on a 3.8 m satellite dish located outside of the cafeteria located in Foothills Lab building 2.

- Setup remote data serving of GOES-16/17 ABI imagery on the top level data servers that we operate in UCAR

  By any measure, the GOES-16/17 imagery is very popular in the community.

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

  This setup will mirror the one that we have operated for GOES-16 GRB products for the past year.

- Since repointing the 4.5m satellite dish at the NCAR Mesa Lab from GOES-16 (75 W to GOES-17 (137.2 W), the Terrestrial Interference (TI) that had been hampering data ingest activities has dropped off to the point of full usability.

  Background:

  In the fall of 2017 we began experiencing significant TI in the GOES-16 signal being received by the 4.5m 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 Foothills Lab 2 (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. As expected, signal levels and quality improve when the leaves fall off of the trees during the fall, and worsen when the leaves return in the spring. 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. In the interim, we have been preparing to install a 3.8m mesh dish (from NCAR/EOL) on the western satellite pad at the Mesa Lab with the expectation that this location will be shielded from the TI problems that plagued our GOES-16 ingest.

Ongoing Activities

We plan to continue the following activities:

- Ingest GOES ReBroadcast (GRB) streams from GOES-16 and GOES-17 in real-time

- Continue to distribute GOES-16 and GOES-17 data via the LDM/IDD and serve the data via the TDS, McIDAS ADDE and AWIPS EDEX

- 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

- Reorganize the contents of the LDM/IDD NIMAGE and UNIWISC datastreams

  GOES ReBroadcast (GRB)

  The UPC has been making all products being distributed in the GRB since installation of downlinks for both GOES-16 (Winter 2017) and GOES-17 (Fall 2018), first as test point-to-point IDD SPARE datastream before the satellites were declared operational, and then in the IDD SATELLITE datastream (formerly known as DIFAX) and via our data server technologies, the THREDDS Data Server (TDS), McIDAS ADDE and AWIPS EDEX.

  The volume of data available in the SATELLITE datastream can be seen in:

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

  NOAAPort SBN

  As part of the transition of GOES-East from GVAR GOES-13 to GOES-R GOES-16, the NWS stopped creating/distributing a wide variety of image sectors, in GINI format, in the NOAAPort SBN. These were replaced by Sectorized Cloud and Moisture Imagery (SCMI) tiles that need to be stitched together to make Full Disk, CONUS and one or two Mesoscale "floater" sectors. As far as we can tell, the NWS' plan is to discontinue the GINI image sectors created from GOES-15 imagery when GOES-15 is turned off in early July. GOES-R GOES-17 assumed the role as GOES-West on February 11.

  The GOES-East/West SCMI tiles that are being distributed in NOAAPort are available (along with some other products) in the IDD NOTHER datastream. The volume of data available in the NOTHER datastream can be seen in:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?NOTHER+oliver.unidata.ucar.edu

  Unidata provides a Python script that is used to stitch together the tiles into full scenes in the ldm-alchemy section of the Unidata GitHub presence, and this utility is being used by several sites.

  IDD NIMAGE Datastream

  Since the process of stitching together NOAAPort-delivered SCMI tiles into usable full scenes requires local processing resources that we believe would be better used for other activities at user sites, we have been working towards distributing reconstituted images created from tiles via the IDD for those sites that would like to continue to receive directly usable images in a manner similar to that for the GINI images that are being discontinued. This is a relatively simple/straightforward effort for the UPC since

we are already stitching together the NOAAPort-delivered tiles and making the reconstituted full scenes available by our TDS, ADDE and EDEX servers. The current plan is to add the reconstituted images to the IDD NIMAGE datastream (NIMAGE was named to represent NOAAPort IMAGE) in the late March/early April time frame.

Also available in the NOAAPort SBN are Level 2 products created from GOES-East imagery. Conversion of the products into forms directly usable by Unidata-supported display and analysis packages is straightforward in that all that needs to be done is to strip off a WMO style header and trailer from the product and then filing the resultant product which is in netCDF4, and the aforementioned ldm-alchemy package contains a Python script that is used to strip the header/trailer. Since it seems logical to provide ready-to-use Level 2 products in the same way that we will provide ready-to-use GOES-East/West imagery, we will add distribution of stripped Level 2 products to the IDD NIMAGE datastream, and this will be done at the same time as for the SCMI imagery.

VALUE-ADDE Products

Texas Tech University (Eric Bruning) is creating value-added Level 2 products created from Geostationary Lightning Mapper (GLM) images as a precursor for similar products potentially being added to NOAAPort. We are currently receiving these products for testing use within the UPC. Our plan is to add these Level 2 products to the NIMAGE IDD datastream when we add the NOAAPort-delivered Level 2 products. The GLM Level 2 products will be directly usable by all of the analysis and display packages that we make available.

IDD UNIWISC Datastream

After we begin making GOES-East/West images reconstituted from NOAAPort tiles and Level 2 products that originate from NOAAPort and value-added providers (we anticipate that TTU will be the first, but not the only site creating value added products), we plan to review the image products that we create for the IDD Unidata-Wisconsin (UNIWISC) datastream and remove those products that are believed to be redundant. Exactly what the removed images will be replaced with will be determined after consultation with the Unidata Governing Committees and with the Unidata community.

- 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

- 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, the TDS, McIDAS ADDE and AWIPS EDEX.

Relevant Metrics

- Lots O Data!

  The volume of GOES-16 and GOES-17 GRB products (12 GB/hour ave and 16 GB/hour max; this is the most voluminous IDD feed!) 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+oliver.unidata.ucar.edu

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

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

  - Groups within UCAR/NCAR (3: all products Unidata, EOL, RAL))
  - U.S. Universities (25: variety of feeds; GLM very popular)
  - U.S. Government (3: all products to 2 NOAA sites and one Military site)
  - International (3: 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.

Status Report: Internet Data Distribution

September 2018 - March 2019

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-16/17 imagery and products being received via redundant GOES-16/17 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-16/17 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:

```
20190301

Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume  94978.758 M bytes/hour
Average hourly volume  63576.489 M bytes/hour

Average products per hour     469487 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)            (M byte/hour)    number/hour
SATELLITE               12243.848     [ 19.258%]    16012.667     6624.489
FSL2                    10456.146     [ 16.447%]    21048.504    11894.447
NGRID                   10052.929     [ 15.812%]    14660.430    70774.000
CONDUIT                  8111.057     [ 12.758%]    21190.362    90587.872
NOTHER                   6179.793     [  9.720%]     9407.057    11261.064
EXP                      5531.266     [  8.700%]     9411.394     5749.936
NEXRAD2                  4569.473     [  7.187%]     5277.099    77155.064
```

FNMOC	3180.255	[5.002%]	8398.569	8665.191
NEXRAD3	1513.229	[2.380%]	1724.658	99316.872
HDS	1144.869	[1.801%]	1510.709	39380.255
GEM	226.633	[0.356%]	1240.950	1407.021
UNIWISC	99.363	[0.156%]	141.946	50.830
FNEXRAD	92.921	[0.146%]	107.701	105.106
NIMAGE	88.933	[0.140%]	136.827	119.511
IDS|DDPLUS	83.771	[0.132%]	104.313	45976.234
LIGHTNING	1.902	[0.003%]	4.882	418.255
GPS	0.100	[0.000%]	1.050	1.021

The 12-16 GB/hr value in IDD data volume shown as SATELLITE represents the ingestion of GOES-16 and GOES-17 data. As expected, the volume is roughly twice what it was when only GOES-16 data was being relayed.

New Data Distribution:

Unidata took 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-test.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
  - HRRRx

- 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: UCAR/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 v6.13.7, but will soon be updated to v6.13.8.

- After discovering that the LDM/IDD Product IDs, which are essentially WMO 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 implemented enhanced processing to insure that all L2 products are processed and then made available via our data services.

Relevant Metrics

- Approximately **505** machines at **234** sites are running LDM-6 **and** reporting real-time statistics to the UPC.

  We know that there are a number of sites that are participating in the IDD, but are not reporting real-time statistics back to us.

  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,

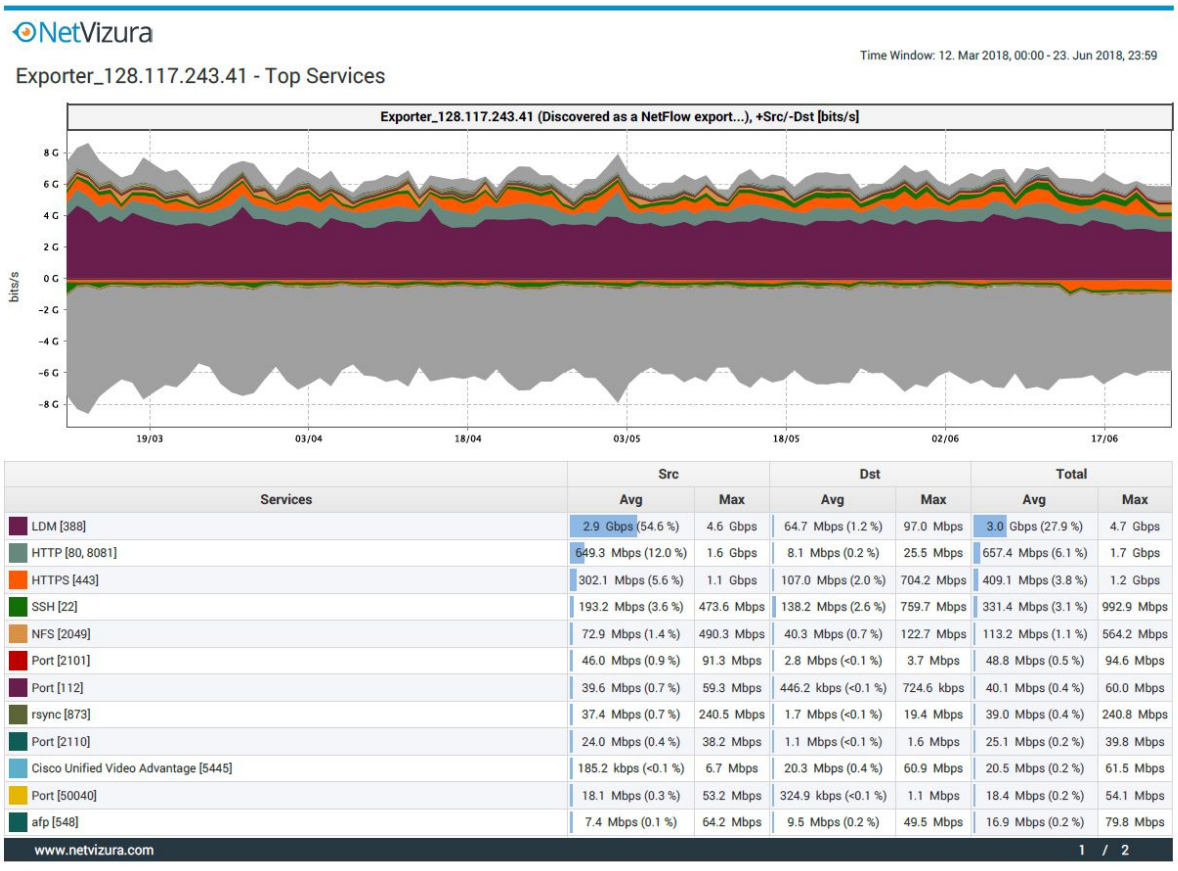
Governments of Spain, South Korea, private companies, etc.).

- 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 1110 downstream connections.

  NB: The following was reported in the Fall 2018 IDD status report. The UCAR site where we generate these statistics has been completely changed, and we have not yet figured out how to get current numbers.

  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

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 *March 2019*

Status Report: IDV with RAMADDA

September 2018 - March 2019

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 4.6.12-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.

__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, Windows app and MacOS certificates have been renewed and will be valid until at least May 30, 2020 (MacOS certificate is valid until 2024). Moreover, as properly signing the IDV under these different environments can be an involved process, this information has thoroughly documented here.

IDV Display Changes

__Integrated DRILSDOWN plugins__

Integrated DRILSDOWN Mapes collection plugin into the IDV release. This includes new display types of probing in pressure coordinates and cross section in pressure coordinates.

__GOES 16/17 GLM ADDE Chooser__

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 a new data chooser to access three GLM source types: Geostationary Lightning Mapper EVENT, GROUP, and FLASH Data through the UNIDATA ADDE data server. The GLM data queries can be performed on the relative times or absolute times.

__3D Streamline Display__

Streamlines are an instantaneous tangent to the velocity vector of the flow and a snapshot of the flow field at some particular time. In the IDV 3D space, streamline display provides the direction in which a massless fluid element will travel at any point in time. 3D streamline display type is available for both derived parameters: Grid 3D Trajectory and 3D Flow Vectors, and the initial area of the 3D streamlines can be a surface or a 3d volume.

__3D Trajectory Display Update__

The grid 3D trajectory display include new display types: Tracer and Tracer Point. The initial area of the 3D Grid Trajectory can be set at XZ or YZ plane, and controlled by the data index in the X or Y direction.

__3D Isentropic Advection__

The new grid 3D isentropic advection parameters interpolate advection from isobaric space into the isentropic space and creates the 3D Theta surface advection displays.

__Grid 3D Vertical Profile in Pressure Axis__

Traditional Skew-T Log-p plots are very poor at understanding tropical convection - moisture interactions, at lower atmosphere. New vertical profile display using pressure as its axis to emphasize information in the lower levels.

__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 helping IDV users to generate WRF-Hydro data in CF compliant Point Feature type data format. In addition, we are helping them with visualization of this dataset in the IDV.

IDV Release

IDV 5.6 will be released in May of 2019.

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.

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

__2019 American Meteorological Conference (AMS) Annual Meeting__

- Visualization of Wind Parcel Trajectories Using McIDAS-V and IDV
- 3D Space Isentropic Analysis in the Unidata's IDV
- VISUALIZATION FOR IMPROVING ALGORITHMIC MEASURES OF EVENTS

__2018/2019 IDV lectures at WRF tutorial workshop__

- Yuan delivered two IDV introduction lectures on two semi-annual WRF tutorial workshops.

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 ~20 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__

https://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 thousands of times.

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.

Prepared March 2019

Status Report: Information Technology

September 2018 - March 2019

Mike Schmidt, Matt Perna

Major Activities

IDD cluster -- We've completed the installation and testing of updated IDD cluster at the NCAR Wyoming SuperComputing Center (NWSC). After NCAR completes their network updates, we expect to cut over to this cluster by the end of March.. Notable changes with this cluster include 10Gb/s network service and the ability for LDM queue sizes to easily double 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. We continue to see a significant signal degradation to complete signal loss when UCAR Facilities operates motorized devices (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. The data reception figures look very good, with very low loss.

Network upgrades -- as UCAR upgrades their backbone infrastructure 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. Two clusters (16 nodes) and two data aggregators are at 10Gb/s service.

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). Unidata staff just moved as a group to use Duo two factor authentication. Initial access to most Unidata and UCAR resources requires some form of two factor authentication.

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.

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 *March 2019*

Status Report: LDM

September 2018 - March 2019

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:

- Extracted GEMPAK parameter tables into their own GitHub repository for ease of updating
- Added command “updategempaktables” to ldmadmin(1)
- Improved responsiveness of LDM system to being stopped
- Improved documentation
- noaaportIngester(1):
    - Improved responsiveness to being externally signaled
    - Improved construction of LDM product-identifier:
        - Enabled creation of a complete LDM product-identifier even if the forecast time-interval of a GRIB2 field can't be decoded due to an unknown PDTN.
        - Added requirement that the GRIB2 version number not be missing (255) and the local version number be 0 or missing (255) in order to obtain the WMO GEMPAK vertical coordinate table.
        - Made the decision to use GEMPAK's WMO parameter table more robust by requiring that the Local Table number be 0 or 255 (missing).
        - Modified how GEMPAK finds a parameter entry:
            - Replaced linear search with binary search
            - Allowed use of parameter with same discipline, category, and ID but different parameter definition template number if an exact match doesn't exist. An ERROR-level message will be logged in this case.
            - Added some error-checking to GRIB2 and GEMPAK libraries
- Eliminated use of portmapper by ldmd(1) because
    - Lots of places consider it a security hole
    - It slowed things down significantly if it wasn't running
    - It doesn't appear to be used much
- Added non-blocking locking of metrics file to ldmadmin(1)’s “addmetrics” command to eliminate concurrent writes by multiple "ldmadmin addmetrics" processes when under severe load.
- Added "-F" option to pqing(1) to enable XON/XOFF software flow control for TTY

input.
- Improved concurrency and thread-safety
- Improved format of log messages
- Added to GEMPAK parameter tables
- Improved code that handles GEMPAK parameter tables
- Modified RPC library to allow interruption by signals
- Repurposed the unused DIFAX feed to SATELLITE for satellite imagery.
- Released versions 6.13.7 and 6.13.8

Dependencies, challenges, problems, and risks include:

- The LDM is sometimes held responsible for decisions made by the NWS on how to categorize data products (not a new challenge)

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.

This project ends April 1st but a no-cost extension has been approved because funding was delayed.

Progress has been made on the following:

- Created OESS-Client module
- Created program for dynamically managing virtual interfaces
- Added support for regular IP multicasting (i.e., multicasting that doesn't use an AL2S multipoint VLAN)
- Added start-of-transmission time to beginning-of-product packet to enable throughput calculation
- Completed integration of multicast capability into LDM
- Added LDM7 server at University of Wisconsin

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 LDM7 servers at more participating universities
- Test LDM7
- 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.

Prepared *March 2019*

Status Report: McIDAS

September 2018 - March 2019

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 2018b was released

  v2018b includes all SSEC versions up to and including the current McIDAS-X and -XCD releases, both of which are v2018.2.

- 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:

- 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, 2019 through February 28, 2019.

  The amount of data served by Unidata ADDE instances increased substantially with the availability of GOES-16/17 ABI imagery and GLM lightning data.

- 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 *March 2019*

Status Report: netCDF

September 2018 - March 2019

Ward Fisher, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Are there other cloud-based block storage formats/locations (zarr, Azure, etc) that are actively in use? That we should consider storing.
2. Are there any emergent avenues (stack overflow, etc) for user support which the netCDF team should investigate?
3. How can we encourage more user testing of the release candidates we provide?

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 **90** open issues for netCDF-C, **43** open issues for netCDF-Fortran, and **15** 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:

- Multiple releases of netCDF-C, netCDF-Fortran.
- Work on the roadmap for different storage formats beyond native netCDF and HDF5-based storage formats.
- Refinement 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-C documentation, modernization of the netCDF-Fortran 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.

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.
- Increase in external contributions has greatly increased the project management overhead for netCDF-C/C++/Fortran.

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++.
- Continue modernizing the netCDF documentation to provide easy access to documentation for older versions of netCDF.
- Publish formal documentation/roadmap for implementing cloud-native storage (object storage).

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 native object storage in the netCDF C library.
- Enhance thread-safety for the netCDF C library.
- Participate in the HDF Technical Advisory Board process.
- Participate in the Zarr/n5 collaboration conference calls.

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 192,041 lines of code (up from 189,774 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 slightly decreased to **0.69,** where it was **0.71** six months ago. 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:

- **857,000** for netCDF-3
- **741,000** for netCDF-4
- **700,000** for HDF5
- **299,000** for GRIB2

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

- **351** for netCDF-3
- **712** for netCDF-4
- **12,000** for HDF5
- **997** 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 2019*

Status Report: Python

September 2018 - March 2019

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 can Unidata's Python training better serve your needs or the needs of your students? Are there other topics we need to add to the workshop? Are there additional opportunities (e.g. conferences) we should explore as convenient venues?
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 roadmap? Anything you notice as lacking in MetPy or Siphon?

Activities Since the Last Status Report

Staffing Changes

John Leeman will unfortunately be stepping down from Unidata in May 2019. A search for a new Python Developer, with an emphasis on instructional skills and experience, has begun. We hope to have the search completed by the end of Spring 2019. Additionally, Ryan May has assumed more responsibilities by joining Unidata's Management Team, which is requiring about 25% of his time. These changes and supporting activities have impacted the bandwidth of the Python development team. This should not have lasting impacts on our capacity, pending successful completion of the hire.

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 creation of training materials, workshop preparation, and logistics against time devoted to support and Python software development. We have also begun looking at ways to unify our various training efforts (gallery, workshop, and online python training) into a proper "Python Training Portal". The hope would be to make it simpler to have community contributions and to turn this into a better resource for the community's Python training needs.

Progress has been made on the following:

- Ryan May and John Leeman, together with Kevin Goebbert, taught a short course

using MetPy at the 2019 AMS Annual Meeting, focused on synoptic meteorology. The course was taught to a sold out audience of 22 attendees. We plan to submit for another course for AMS 2020.
- Ryan May, Sean Arms, and Howard Van Dam taught a 1-day workshop at Metro State University in Denver in March to 23 attendees.
- Ryan May and John Leeman taught a regional workshop at Valparaiso University on 14-15 March. This workshop also included a Saturday morning "Hack Day".
- Ryan may and John Leeman will teach a workshop on testing in Python at the NCAR-hosted Conference on Improving Scientific Software in April 2019.
- 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. Unidata's YouTube channel has had 49.6k minutes of watch time in the last year (up 108% from the previous year), 17.4k views (up 110%), 242 new subscribers (up 105%), and the most popular video has received over 2000 views. We welcome additional community screencasts and suggested topics.

MetPy

MetPy continues to grow, both in features and in community. The volume of support requests is the most remarkable area of growth; traffic and activity across GitHub, E-Support/E-mail, and Stack Overflow is steady. Community code contributions have been somewhat slower; this may be due in part to less activate solicitation of contributions, owing to the uptick in support requests and aforementioned staffing changes.

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 the GEMPAK-like interface, improved units support, integration with xarray, and data formats. We do anticipate the release of MetPy 1.0 this year. Also, to try to foster more community discussion of MetPy's goals and plans, we have published a general MetPy roadmap that tries to capture our plans from GitHub in a more friendly format.

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 749 followers (29% growth in 6 months).
- MetPy 0.9.2 was released with a few minor bug fixes
- MetPy 0.10 was released, including the initial GEMPAK-like plotting functionality, more xarray integration, and various calculation enhancements.
- 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. The most pressing developments we anticipate for Siphon are improvements to working with Siphon in interactive sessions, like the Jupyter notebook environment: improved catalog crawling interface, better string

representations, and tab completion. Siphon continues to see community contributions trickle in. We hope to have one of Unidata's summer interns do some contributions to Siphon as part of their summer work.

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. We have seen several community questions regarding both the GOES and NEXRAD processing software.

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. Ryan May continues to serve as a core developer for CartoPy as well as a member of Matplotlib's Steering Council.

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.
- 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. This engagement is enhanced by work on the Pangeo EarthCube award, which will likely drive some contributions to the XArray project
- Ryan May served as the release manager for the CartoPy 0.17 release
- We also continue to actively engage with the xarray, numpy, and pint projects

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.

Changes to python-awips since the last report include (through release 18.1.7):

- New functions **DataAccessLayer.getMetarObs()** and **DataAccessLayer.getSynopticObs()** added to process retrieved surface parameters into a dictionary.

- Added a new class **awips.dataaccess.ModelSounding()** to request vertical soundings from any available AWIPS model with isobaric data levels.
- New methods added to **DataAccessLayer** called **getRadarProductNames()** and **getRadarProductIDs()** to return either names or numerical IDs from the list of products available for the radar datatype.
- Added GEMPAK-specific scripts for processing data from EDEX to GEMPAK/NMAP2 display.
- Better control for Python 3 bytestring encoding Python 2 unicode.
- New Jupyter Notebooks using python-awips:
    - GLM Point Data
    - Profiler Time-series
    - METAR Station Plot
    - Colored Surface Temperature
    - Regional METAR/SYNOP Surface Obs using geographic boundaries from the AWIPS database (no cartopy or basemap required)
    - Forecast Model Vertical Sounding

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:

- Teach Python workshop for Unidata/UCAR/NCAR interns
- Teach additional Python regional workshops
- Attend SciPy 2019

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

- Teach another short course on MetPy at AMS 2020
- Present annual update on Python libraries at AMS 2020

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 training materials into a more unified Python training portal

Relevant Metrics

MetPy

- 98% test coverage
- Watchers: 51
- Downloads for the releases made in the last year (only Conda for now):
    - 0.8.0: 5314
    - 0.9.0: Not released
    - 0.9.1: 3216
    - 0.9.2: 8405
    - 0.10.0: 2661
- Since 1 September 2018
    - Active Issues: 74 (43 created, 21 closed)
    - Active PRs: 38 (33 created, 30 closed)
    - External Issue Activity: 24 opened, 50 comments
    - External PR Activity: 14 opened, 9 comments
    - Unique external contributors: 31
    - Stars: 76 (380 total)
    - Forks: 0 (133 total)
    - Commits: 93
- Since 1 March 2018
    - Active Issues: 202 (126 created, 100 closed)
    - Active PRs: 149 (126 created, 121 closed)
    - External Issue Activity: 63 opened, 194 comments
    - External PR Activity: 35 opened, 69 comments
    - Unique external contributors: 62
    - Stars: 145 (380 total)
    - Forks: 3 (156 total)
    - Commits: 384

Siphon

- 97% test coverage
- Watchers: 14
- Downloads for the last year (only Conda for now):
    - 0.7.0: 9306
    - 0.8.0: 10889
- Since 1 September 2018:
    - Active Issues: 17 (14 created, 1 closed)
    - Active PRs: 7 (5 created, 5 closed)
    - External Issue Activity: 7 opened, 13 comments

    - External PR Activity: 2 opened, 0 comments
    - Unique external contributors: 8
    - Stars: 11 (90 total)
    - Forks: 1 (35 total)
    - Commits: 12
- Since 1 March 2018
    - Active Issues: 55 (37 created, 30 closed)
    - Active PRs: 39 (38 created, 37 closed)
    - External Issue Activity: 16 opened, 40 comments
    - External PR Activity: 7 opened, 8 comments
    - Unique external contributors: 22
    - Stars: 29 (90 total)
    - Forks: 1 (35 total)
    - Commits: 133

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

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 *March 2019*

Status Report: Support

September 2018 - March 2019

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

- Starting in 2018, the UPC has been focusing its in-person training efforts on regional workshops and short courses.
- 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

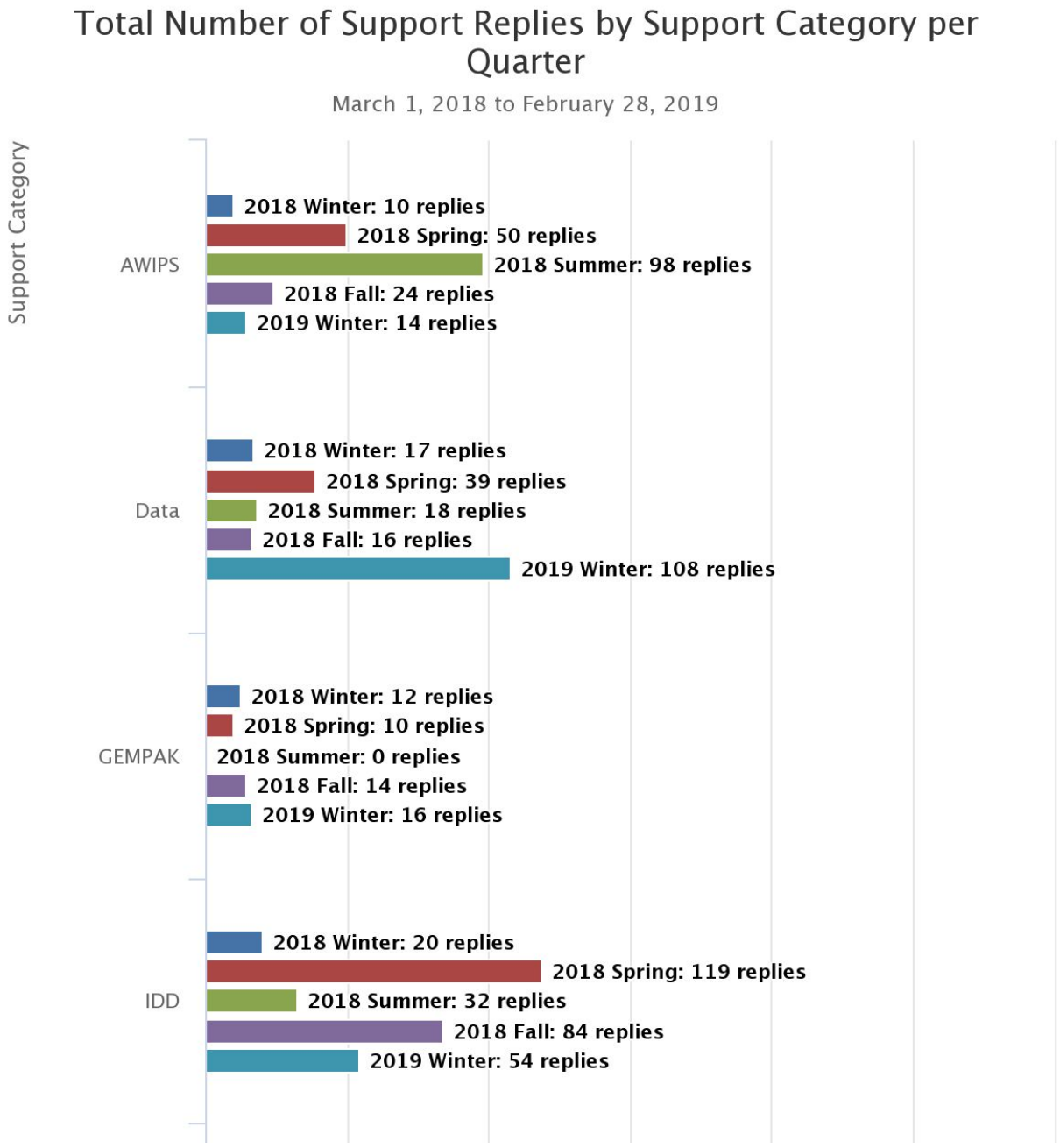
Since January 26, 2006 over 57950 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

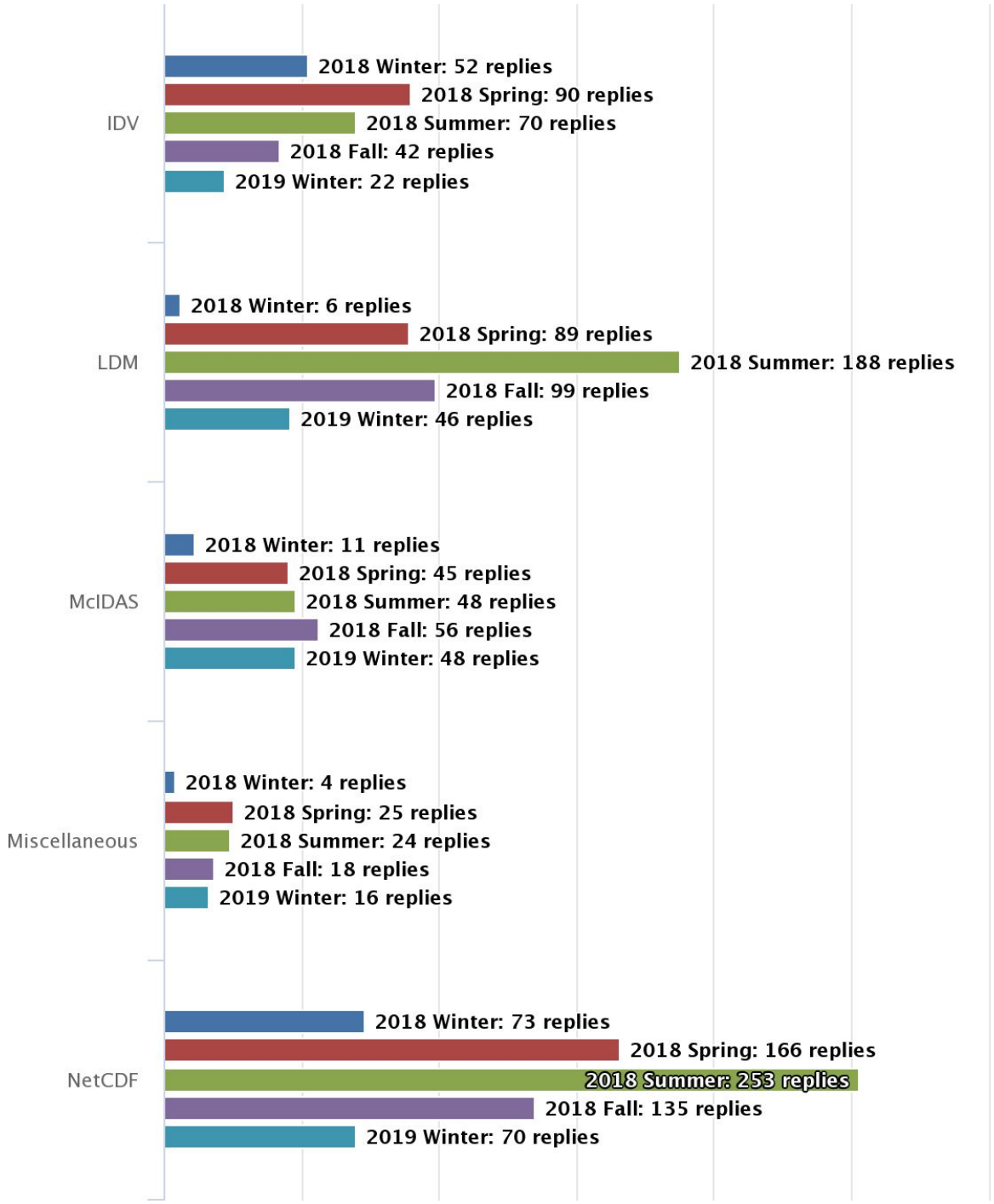
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

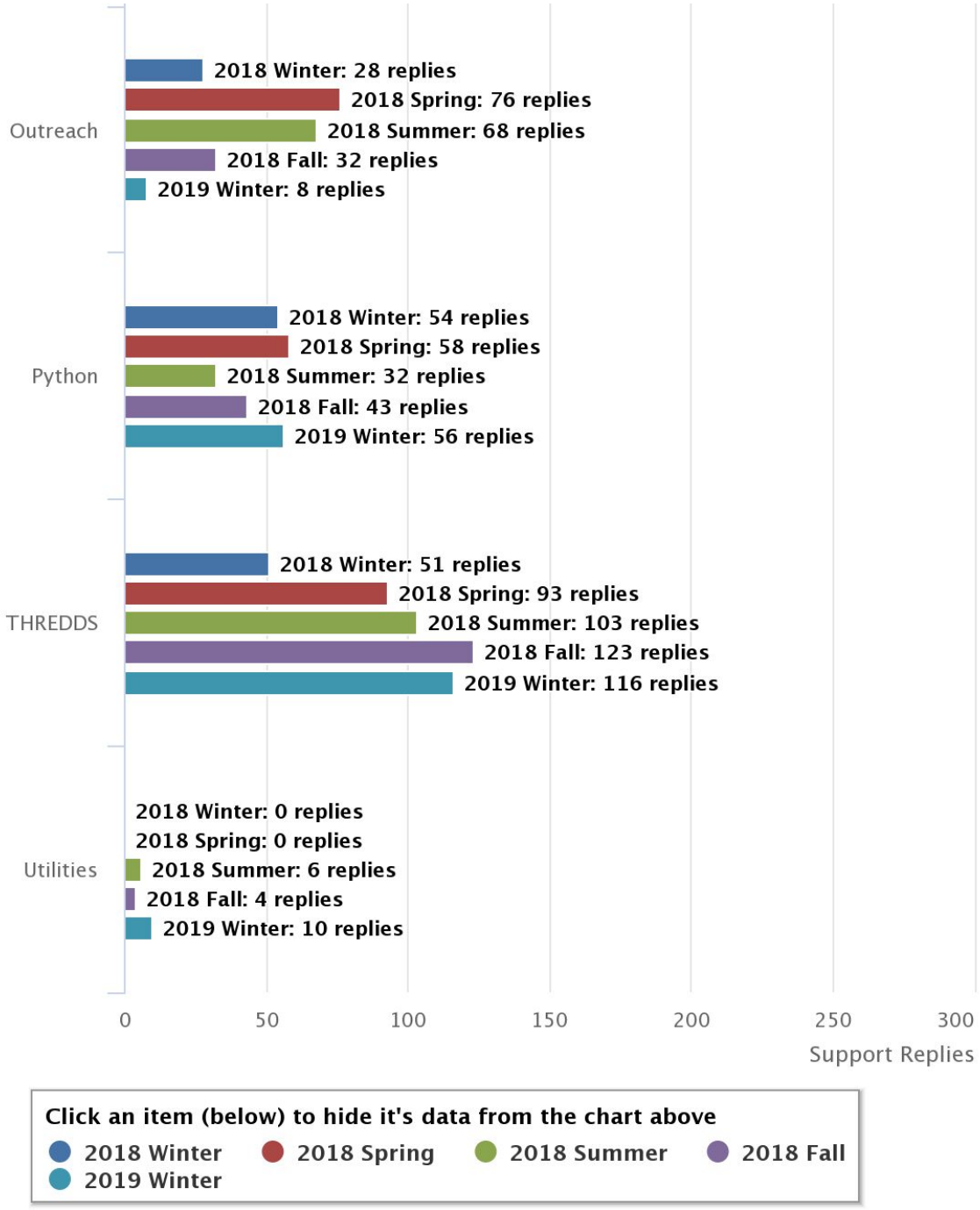
Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 14 month period from March 1, 2018 until February 28, 2019.

The quarters shown are defined as:

Winter:	**Spring:**	**Summer:**	**Fall:**
January, February, March	April, May, June	July, August, September	October, November, December




IDV
2018 Winter: 52 replies
2018 Spring: 90 replies
2018 Summer: 70 replies
2018 Fall: 42 replies
2019 Winter: 22 replies
LDM
2018 Winter: 6 replies
2018 Spring: 89 replies
2018 Summer: 188 replies
2018 Fall: 99 replies
2019 Winter: 46 replies
McIDAS
2018 Winter: 11 replies
2018 Spring: 45 replies
2018 Summer: 48 replies
2018 Fall: 56 replies
2019 Winter: 48 replies
Miscellaneous
2018 Winter: 4 replies
2018 Spring: 25 replies
2018 Summer: 24 replies
2018 Fall: 18 replies
2019 Winter: 16 replies
NetCDF
2018 Winter: 73 replies
2018 Spring: 166 replies
2018 Summer: 253 replies
2018 Fall: 135 replies
2019 Winter: 70 replies


Outreach
2018 Winter: 28 replies
2018 Spring: 76 replies
2018 Summer: 68 replies
2018 Fall: 32 replies
2019 Winter: 8 replies
Python
2018 Winter: 54 replies
2018 Spring: 58 replies
2018 Summer: 32 replies
2018 Fall: 43 replies
2019 Winter: 56 replies
THREDDS
2018 Winter: 51 replies
2018 Spring: 93 replies
2018 Summer: 103 replies
2018 Fall: 123 replies
2019 Winter: 116 replies
Utilities
2018 Winter: 0 replies
2018 Spring: 0 replies
2018 Summer: 6 replies
2018 Fall: 4 replies
2019 Winter: 10 replies
0
50
100
150
200
250
300
Support Replies
Click an item (below) to hide it's data from the chart above
2018 Winter
2018 Spring
2018 Summer
2018 Fall
2019 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, Student Interns
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 shown a slight decline over the past several years, but this is most likely attributable by the increased ways support is being provided. 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.
- 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 several 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 *March 2019*

Status Report: THREDDS

September 2018 - March 2019

Sean Arms, Dennis Heimbigner, Howard Van Dam II

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Please help us welcome Howard Van Dam to Unidata and the THREDDS team!
2. Is this useful? (see the metrics section for context and details)

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 4.6.13 (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, bug, and security fixes, continues to take up already severely limited resources.
- Version 4.6.13 was released on February 19th, 2019. Version 4.6.12 was ready in December; the choice was made to not announce its availability due to the government shutdown because of the security vulnerabilities addressed by the 3rd party library upgrades.
- 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 beta6 of the TDS. Please check out the new beta on our thredds-test server; we anticipate a release candidate for TDS 5.0 in the summer of 2019.

Progress has been made on the following:

- The TDS now has experimental support for generating Dataset elements from schema.org. This allows datasets served by the TDS to be picked up by Google’s

Dataset Search (example). This is completely opt-in, and requires a few extra steps by server admins to submit their datasets for inclusion.

- Support the netCDF-CF simple-geometry spec within the CDM (Contribution by staff members at USGS funded by a Water Resources Research Institute)
- Our new documentation has made great strides as part of the preparation for the Fall 2018 workshop.
- Thanks to our Summer Intern, 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.
    - Presented results at the 2018 Fall AGU meeting and at AMS 2019.
    - These are being actively used by a few of those running the latest beta version of the TDS.
- 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.
    - Our last updated indicated that we are struggling to keep the ansible build working as it is such a radical departure from the norm. We have since abandoned this work and will be managing Nexus in a more traditional way.
- 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.
- A very basic, proof of concept Upload/Download support has been added to TDS. This now includes an upload web form accessible as http://.../thredds/upload.
- ncSOS has been integrated into the TDS distribution (as part of the OIIP project—see the Rosetta section for more details). Applied Science Associates is looking into porting this to v5.0.
- 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:

- While all java based components in the THREDDS project run under Java 11, we are currently building using Java 8. Portions of our build infrastructure need to be reworked to use Java 11; this was one of Christian's areas of expertise, and it taking quite some time to get addressed.
- 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, Peter Pokrant, Victor Gensini, the NCAR RDA, etc.!).

Rosetta

Rosetta continues to progress following a very successful 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. A poster was presented at the Fall AGU 2018 meeting with respect to the advances in Rosetta related to OIIP. We continue to work with JPL as part of their user acceptance process, with the end goal being the operational use of Rosetta at the PO.DAAC.

Progress has been made on the following:

- Cleaning up the UI based on user feedback
    - Created a unified workflow for the gui wizard interface (one workflow for the three major CF DSGs [timeseries, profile, trajectory]).
    - Front-end driven by the JPL developed Metadata Profile Service, which ensures updated and accurate collection of required/recommended/optional metadata
- The entire front end has been reworked, and no longer depends on the abandoned jWizard library. As such, our jQuery lib is up-to-date.
- Several youtube videos available on the OIIP YouTube channel demonstrating the interactive and batch processing workflows of Rosetta.

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. As part of the preparation for our Fall 2018 workshop, the

second pass at the overhaul is now complete.
- Maintain thredds.ucar.edu and keep up with the addition of new datasets to the IDD.
    - Note that the TDS happily works with FV3 GFS output, and will be made available as soon as it ships across NOAAPORT and/or IDD.
- 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.
- We are partnering with JPL / PO.DAAC on a ASA ROSES Advanced Information Systems Technology (AIST) solicitation. The proposal is tentatively titled “MIKITA – Multi-sensor Data Integration Microservices for Knowledge InTensive Applications.” The focus of the work is on extending the Metadata Profile Service (MPS) and creating Semantic Technology based microservices which leverage the MPS. WHile the bulk of the work is on the JPL side, we are proposing to extend Rosetta to interact directly with the MPS through its web API, and extending the TDS to provide metadata records in a more semantic friendly format, such as JSON-LD based on schema.org. While we are Unidata are not experts on the semantic web, we can certainly empower those who are.
- 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:

- 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 (release candidate targeted for Summer 2019).

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

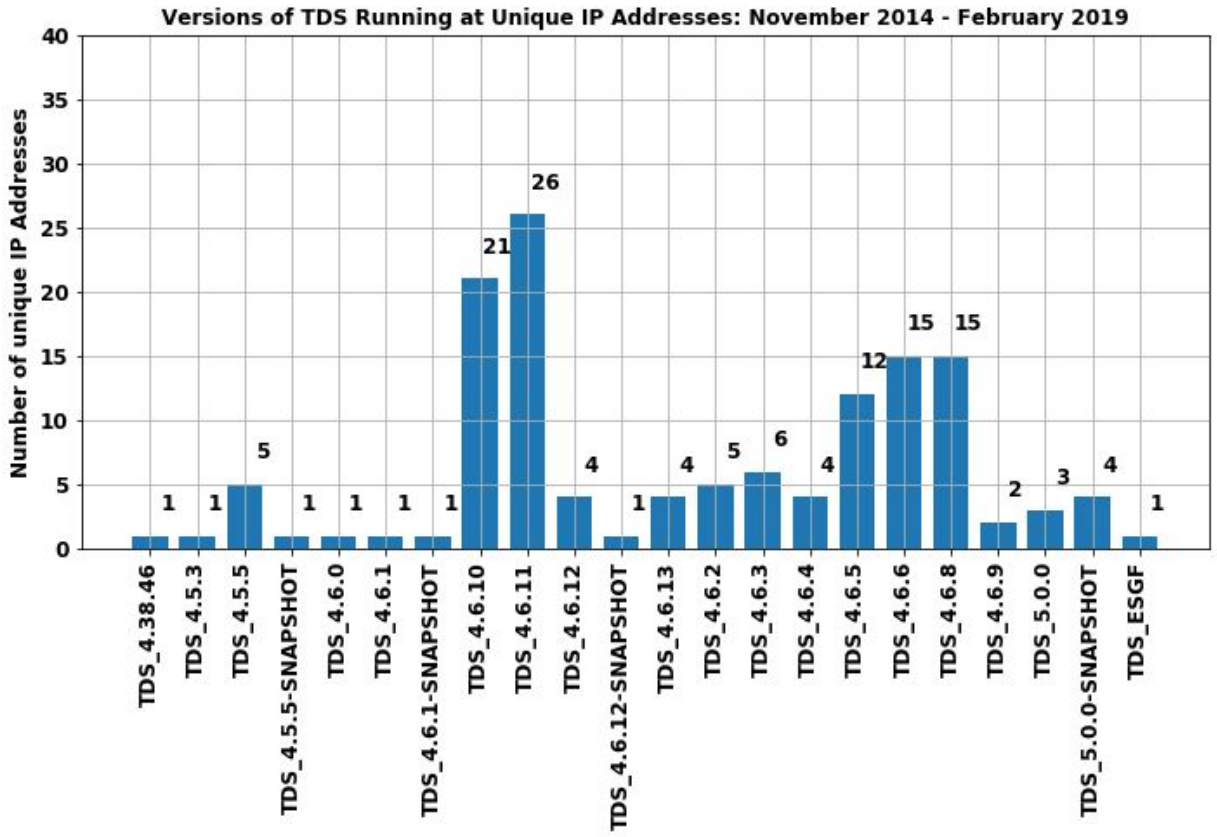
- Support Zarr and, potentially, TileDB at the IOSP layer of the CDM.
- Support the UGRID spec within the CDM
- Complete NASA JPL user acceptance testing for Rosetta
- Create a TDS plugin layer for external services (potentially use Project Jigsaw from Java 9)

    - Upgrade the ncWMS, ncISO, and other plugin services to use the new TDS 5.x plugin layer
- 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.
- 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



19,155 unique IPs started up thredds from November 2014 through February 2019, **134** 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

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. This can also happen if the TDS has been configured to use a different context without the use of a proxy server.
- 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 (v4.6.13 was the most current release during this period, and was released on 12 February 2019).

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.

Furthermore, of the **134** publically accessible servers, **74** have updated the name of their server in their server configuration file (taken as a sign that they are maybe possibly intended to be used by others...maybe…). A list of those servers can be found here.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   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 *March 2019*